

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 STQTPNG 8
|:|||||
Db 1023 SNQSPNG 1030

RESULT 15
O95875 PRELIMINARY; PRT; 2157 AA.
AC O95875;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE BAT2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Rowen L., Madan A., Qin S., Shaffer T., James R., Ratcliffe A.,
RA Abbasi N., Dickhoff R., Loretz C., Madan A., Dors M., Young J.,
RA Lasky S., Hood L.;
RT "Sequence of the human major histocompatibility complex class III
region.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR ENBL; AFI29756; AAD18086.1; -;
SQ SEQUENCE 2157 AA; 228869 MW; 21B817F5B699B0DB CRC64;

Query Match 72.0%; Score 36; DB 4; Length 2157;
Best Local Similarity 66.7%; Pred. No. 5.9e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 STQTPNGR 9
|:|||||
Db 2073 SSRTPTGR 2081

Search completed: December 11, 2003, 18:28:33
Job time : 24.3333 secs

Q9D582;
01-JUN-2001 (TRENBLrel. 17, Created)
01-JUN-2001 (TRENBLrel. 17, Last sequence update)
01-MAR-2003 (TRENBLrel. 23, Last annotation update)
4921530G03Rik protein.
MMRN OR 4921530G03Rik.
Mus musculus (Mouse).
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Azawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Giesi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Stauber F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojuna N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombarts P.,
RA Nordone P., Ring B., Schenwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.,
RA "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
DR EMBL; AK014984; BAB29654.1; -
DR MGD; MGI:1918195; Wmrn.
SQ SEQUENCE 193 AA; 20850 MW; 14E1207E909DCBA2 CRC64;

Query Match 72.0%; Score 36; DB 11; Length 193;
Best Local Similarity 75.0%; Pred. No. 49;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 TOTPPNGR 9
DB 88 TSTPPDGR 95

RESULT 10
Q38155 PRELIMINARY; PRT; 265 AA.
AC Q38155;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Gp37, tail fiber (Fragment).
OS Bacteriophage SV14.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.
ON NCBI_TaxID=47488;
RN [1]
RP SEQUENCE FROM N.A.
RA Tetart F., Repolia F., Monod C., Krisch H.;
RT "Bacteriophage T4 host range is expanded by duplications of a small
RL domain of the tail fiber adhesin,";
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z67978; CAA91919.1; -
DR InterPro; IPR005003; Phage fiber.
DR Pfam; PF03335; Phage fiber; 6.
FT NON_TER 1
SQ SEQUENCE 265 AA; 27435 MW; 62FB6B9B5F9124E5 CRC64;

Query Match 72.0%; Score 36; DB 9; Length 265;
Best Local Similarity 85.7%; Pred. No. 67;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TOTPPNGR 8
DB 25 TDTPPNG 31

RESULT 11
Q8YS06 PRELIMINARY; PRT; 290 AA.
AC Q8YS06;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Hypothetical protein A113286.
GN A113286.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
ON NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2159285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Matanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
cyanobacterium Anabaena sp. strain PCC 7120."
RL DNA Res. 8:205-213(2001).
DR EMBL; AP003592; BAB74985.1; -
DR InterPro; IPR002502; Amidase 2.
DR Pfam; PF01510; Amidase 2; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 290 AA; 32444 MW; 4775FB9679478899 CRC64;

Query Match 72.0%; Score 36; DB 16; Length 290;
Best Local Similarity 66.7%; Pred. No. 74;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 STOTPPNGR 9
DB 208 SLETPDGR 216

RESULT 12
Q99362 PRELIMINARY; PRT; 382 AA.
ID Q99362;
AC Q99362;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Protein 37 (Fragment).
GN 37.
OS Bacteriophage T4.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
OC T4-like viruses.
ON NCBI_TaxID=10865;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T4;
RX MEDLINE=91073397; PubMed=2147721;
RA Montag D., Haselholtsseini S., Henning U.;
RT "Receptor-recognizing proteins of T-even type bacteriophages The
RL receptor-recognizing area of proteins 37 of phages T4 Tu1a and Tu1b."
RL J. Mol. Biol. 216:327-334(1990).
DR EMBL; X55190; CAA38973.1; -
DR InterPro; IPR005003; Phage fiber.
DR Pfam; PF03335; Phage fiber; 7.
FT NON_TER 1
SQ SEQUENCE 382 AA; 39950 MW; D2BE4C650AB7719E CRC64;

Query Match 72.0%; Score 36; DB 9; Length 382;
Best Local Similarity 85.7%; Pred. No. 98;

Query Match 78.0%; Score 39; DB 16; Length 642;
Best Local Similarity 87.5%; Pred. No. 46;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TOTPPNGR 9
DB 408 TOTPPNDR 415

RESULT 2
QSP992 PRELIMINARY; PRT; 642 AA.
AC QSP992;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein XCC1973.
GN XCC1973.
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=340;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Rainach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Catarotte G., Canavan F., Cardozo J., Chamberg F., Ciapina L.F.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Rossi M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Nova J.A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spindola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Stutabai J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463(2002).
DR EMBL: AE012301; ANA41262.1; -
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 642 AA; 70969 MW; C4F639CD10B9C590 CRC64;

Query Match 78.0%; Score 39; DB 16; Length 642;
Best Local Similarity 87.5%; Pred. No. 46;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TOTPPNGR 9
DB 408 TOTPPNDR 415

RESULT 3
Q9RU62 PRELIMINARY; PRT; 1054 AA.
AC Q9RU62;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Transcription-repair coupling factor.
GN DR1532.
OS Deinococcus radiodurans.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1;

Query Match 76.0%; Score 38; DB 16; Length 1054;
Best Local Similarity 77.8%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 STQTPPNGR 9
DB 689 SIQTPPKGR 697

RESULT 4
Q923A9 PRELIMINARY; PRT; 950 AA.
AC Q923A9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Similar to RIKEN GNA 3110039B05 gene (Fragment).
GN BAT2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC006664; AAH06664.1; -
DR MGD; MGI:1915467; Bat2.
FT NON TER 1
SQ SEQUENCE 950 AA; 99726 MW; E41DE9AB4A66D976 CRC64;

Query Match 74.0%; Score 37; DB 11; Length 950;
Best Local Similarity 66.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 STQTPPNGR 9
DB 866 SSRIPPSGR 874

RESULT 5
Q9C2H8 PRELIMINARY; PRT; 976 AA.
ID Q9C2H8;
AC Q9C2H8;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

Query Match 74.0%; Score 37; DB 11; Length 950;
Best Local Similarity 66.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 STQTPPNGR 9
DB 866 SSRIPPSGR 874

RESULT 5
Q9C2H8 PRELIMINARY; PRT; 976 AA.
ID Q9C2H8;
AC Q9C2H8;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

ACT_SITE	130	130	MODIFICATION RESULTS IN INHIBITION OF ATP SYNTHESIS.
DISULFID	240	246	
SEQUENCE	364 AA;	40074 MW;	94F0FA9B444B3EEA CRC64;
Query Match	68.0%;	Score 34;	DB 1; Length 364;
Best Local Similarity	85.7%;	Fred.No. 50;	
Matches	6;	Conservative	0; Mismatches 1; Indels 0; Gaps 0;
1 STQTPPN 7			
11111111			
22 STQAPPN 28			

RESULT 15			
PKNA MYCTU	STANDARD;	PRT;	431 AA.
717585;			
30-MAY-2000	(Rel. 39, Created)		
30-MAY-2000	(Rel. 39, Last sequence update)		
15-SEP-2003	(Rel. 42, Last annotation update)		
Probable serine/threonine-protein kinase pKNA	(EC 2.7.1.37).		
PKNA OR RV0015C OR MT0018 OR MTCY10H4.15C.			
Mycobacterium tuberculosis.			
OS			
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;		
OC	Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.		
OC	NCB_TaxID=1773;		
OX	[1]		
RP	SEQUENCE FROM N.A.		
RP	STRAIN=H37RV;		
RC	MEDLINE=98295987; PubMed=9634230;		
RC	Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,		
RA	Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaiia F.,		
RA	Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,		
RA	Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holtroyd S.,		
RA	Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,		
RA	Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,		
RA	Rutter S., Seeger K., Skelton S., Squares S., Squares R.,		
RA	Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;		
RA	"Deciphering the biology of Mycobacterium tuberculosis from the		
RT	complete genome sequence."		
RL	Nature 393:537-544 (1998).		
RL	[2]		
RP	SEQUENCE FROM N.A.		
RP	STRAIN=CDC 1551 / Oshkosh;		
RA	Fleischmann J., Alland D., Eisen J.A., Carpenter L., White O.,		
RA	Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,		
RA	Kolony J.F., Nelson W.C., Unayam L.A., Ermolaeva M.D., Salzberg S.L.,		
RA	Delcher A., Utterback T., Weldman J., Khouri H., Gill J., Mikula A.,		
RA	Bisai W.;		
RA	"Whole genome comparison of Mycobacterium tuberculosis clinical and		
RT	laboratory strains."		
RL	Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.		
CC	-1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.		
CC	-1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announcement/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL; Z80233; CAB02435.1; --		
DR	EMBL; AE006916; AAK44240.1; --		
DR	PIR; E70699; E70699.		
DR	HSP; Q06486; ICK1.		
DR	TIGR; MT0018; --		
DR	TubercuList; RV0015c; --		
DR	InterPro; IPR000719; Prot kinase.		
DR	InterPro; IPR002290; Ser Thr_pkinase.		
DR	Pfam; PF00069; pkinase; 1.		

DR InterPro: IPR001323; EPO_TPO.
 DR InterPro: IPR003978; thrombopoietin.
 DR Pfam: PF00758; EPO_TPO: 1.
 DR PRINTS: PR01485; THROMBOPTN.
 DR PROSITE: PS00817; EPO_TPO: 1.
 KW Cytokine; Glycoprotein; Hormone; Signal.
 FT SIGNAL 1 23
 FT CHAIN 24 352 THROMBOPOIETIN.
 FT FT 24 352
 FT DISULFID 28 172 POTENTIAL.
 FT FT 28 172
 FT DISULFID 50 106 POTENTIAL.
 FT FT 50 106
 FT CARBOHYD 185 165 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT FT 185 165
 FT CARBOHYD 197 197 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT FT 197 197
 FT CARBOHYD 206 206 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT FT 206 206
 FT CARBOHYD 234 234 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT FT 234 234
 FT CARBOHYD 255 255 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT FT 255 255
 FT CARBOHYD 332 332 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT FT 332 332
 FT CARBOHYD 347 347 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT FT 347 347
 SQ SEQUENCE 352 AA; 37641 MW; 024F3B41B061FBD8 CRC64;
 Query Match 68.0%; Score 34; DB 1; Length 352;
 Best Local Similarity 75.0%; Pred. NO. 49;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 2 TOTPPNGR 9
 DB 131 TQLPPQGR 138
 RESULT 13
 TPO_HUMAN
 ID_TPO_HUMAN STANDARD; PRT; 353 AA.
 AC P40225; Q13020; Q15790; Q15791; Q15792;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Thrombopoietin precursor (Megakaryocyte colony stimulating factor)
 DE (Myeloproliferative leukemia virus oncogene ligand) (C-mpl ligand)
 DE (ML) (Megakaryocyte growth and development factor) (MGDF).
 GN TPO.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OK NCBI_Taxid=9606;
 [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Fetal liver;
 RX MEDLINE=94261202; PubMed=8202154;
 RA de Sauvage F.J., Haas P.E., Spencer S.D., Malloy B.E., Gurney A.L.,
 RA Spencer S.A., Darbonne W.C., Henzel W.J., Wong S.C., Kuang W.-J.,
 RA Oles K.J., Hultgren B., Solberg L.A. Jr., Goeddel D.V., Eaton D.L.;
 RT "Stimulation of megakaryocytopoiesis and thrombopoiesis by the c-Mpl
 RT ligand.";
 RL Nature 369:533-538(1994).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Fetal liver;
 RX MEDLINE=94291201; PubMed=8020099;
 RA Bartley T.D., Bogenberger J., Hunt P., Li Y.-S., Lu H.S., Martin F.,
 RA Chang M.-S., Samal B.B., Nichol J.L., Swift S., Johnson M.J.,
 RA Hsu R.-Y., Parker V.P., Suggs S., Skrine J.D., Merewether L.A.,
 RA Clogson C., Hsu E., Hokom M.M., Hornkohl A., Choi E., Pangellinan M.,
 RA Sun Y., Mar V., McNich J., Simcnet L., Jacobsen F., Xie C.,
 RA Shutter J., Chute H., Basu R., Selander L., Trollinger D., Sieu L.,
 RA Padilla D., Trail G., Elliott G., Izumi R., Covey T., Crouse J.,
 RA Garcia A., Xu W.-del Castillo J., Biron J., Cole S., Hu M.C.-T.,
 RA Pacifici R., Ponting I., Saris C., Wen D., Yung Y.P., Lin H.,
 RA Bosselman K.A.;
 RT "Identification and cloning of a megakaryocyte growth and development
 RT factor that is a ligand for the cytokine receptor Mpl.";
 RL Cell 77:1117-1124(1994).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=95109091; PubMed=7809166;

RA Foster D.C., Sprecher C.A., Grant F.J., Kramer J.M., Kuipper J.L.,
 RA Holly R.D., Whitmore T.E., Heipel M.D., Bell L.A.N., Ching A.F.,
 RA McGrane V., Hart C., O'Hara P.J., Lok S.;
 RT "Human thrombopoietin: gene structure, cDNA sequence, expression, and
 RT chromosomal localization.";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:13023-13027(1994).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=95010765; PubMed=7926023;
 RA Schma Y., Akahori H., Seki N., Hori T.-A., Ogami K., Kawamura K.,
 RA Miyazaki H.;
 RT "Molecular cloning and chromosomal localization of the human
 RT thrombopoietin gene.";
 RL FEBS Lett. 353:57-61(1994).
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RX MEDLINE=95152076; PubMed=7849319;
 RA Gurney A.L., Kuang W.-J., Xie M.-H., Malloy B.E., Eaton D.L.,
 RA de Sauvage F.J.;
 RT "Genomic structure, chromosomal localization, and conserved
 RT alternative splice forms of thrombopoietin.";
 RL Blood 85:981-988(1995).
 RN [6]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX TISSUE=Liver;
 RC MEDLINE=96015174; PubMed=8537317;
 RA Kato T., Ogami K., Shimada Y., Iwamatsu A., Schma Y., Akahori H.,
 RA Horie K., Kokubo A., Kudo Y., Maeda E., Kobayashi K., Ohashi H.,
 RA Ozawa T., Inoue H., Kawamura K., Miyazaki H.;
 RT "Purification and characterization of thrombopoietin.";
 RL J. Biochem. 118:229-236(1995).
 RN [7]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX TISSUE=Placenta;
 RC MEDLINE=95122483; PubMed=7822271;
 RA Chang M., McVinch J., Basu R., Shutter J., Hsu R., Perkins C., Mar V.,
 RA Suggs S., Welcher A., Li L., Lu H., Bartley T., Hunt P., Martin F.,
 RA Samal B., Bogenberger J.;
 RT "Cloning and characterization of the human megakaryocyte growth and
 RT development factor (MGDF) gene.";
 RL J. Biol. Chem. 270:511-514(1995).
 RN [8]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 3).
 RA Im S.H., Lee W.S., Chung K.H.;
 RT "Cloning and sequencing of human thrombopoietin.";
 RL Submitted (May-1996) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: LINEAGE-SPECIFIC CYTOKINE AFFECTING THE PROLIFERATION
 CC AND MATURATION OF MEGAKARYOCYTES FROM THEIR COMMITTED PROGENITOR
 CC CELLS. IT ACTS AT A LATE STAGE OF MEGAKARYOCYTE DEVELOPMENT. IT
 CC MAY BE THE MAJOR PHYSIOLOGICAL REGULATOR OF CIRCULATING PLATELETS.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=3;
 CC Name=1;
 CC IsoId=P40225-1; Sequence=Displayed;
 CC Name=2; Synonyms=TPO-2;
 CC IsoId=P40225-2; Sequence=VSP_001450;
 CC Name=3; Synonyms=Truncated;
 CC IsoId=P40225-3; Sequence=VSP_001451;
 CC -!- DOMAIN: TWO-DOMAIN STRUCTURE WITH AN ERYTHROPOIETIN-LIKE N-
 CC TERMINAL AND A SER/PRO/THR-RICH C-TERMINAL.
 CC -!- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
 CC -!- DATABASE: NAME=Rad Systems' cytokine source book: TPO;
 CC WWW="http://www.rndsystems.com/asp/g/sitebuilder.asp?bodyId=225".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----


```

DR EMBL; U87449; AAB47565.1; -.
DR PIR; A55962; A55962.
DR HSP; P02699; IBOU.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR InterPro; IPR001760; Opsin.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHOOPS.
DR PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS0262; G-PROTEIN_RECEP_F1_2; 1.
DR PROSITE; PS00238; OPSIN; 1.
KW Phosphoreceptor; Retinal protein; Transmembrane; Glycoprotein;
KW Phosphorylation; Lipoprotein; Palmitate; G-protein coupled receptor.
FT DOMAIN 1 30 EXTRACELLULAR.
FT DOMAIN 31 55 1 (POTENTIAL).
FT DOMAIN 56 67 CYTOPLASMIC.
FT DOMAIN 68 92 2 (POTENTIAL).
FT TRANSMEM 93 107 EXTRACELLULAR.
FT TRANSMEM 108 127 3 (POTENTIAL).
FT TRANSMEM 128 146 CYTOPLASMIC.
FT TRANSMEM 147 170 4 (POTENTIAL).
FT TRANSMEM 171 194 EXTRACELLULAR.
FT TRANSMEM 195 222 5 (POTENTIAL).
FT TRANSMEM 223 244 CYTOPLASMIC.
FT TRANSMEM 245 268 6 (POTENTIAL).
FT TRANSMEM 269 296 EXTRACELLULAR.
FT TRANSMEM 277 301 7 (POTENTIAL).
FT TRANSMEM 302 351 CYTOPLASMIC.
FT CARBOHYD 4 4 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 192 192 N-LINKED (GLCNAC. .) (POTENTIAL).
FT DISULFID 104 181 BY SIMILARITY.
FT BINDING 288 288 RETINAL CHROMOPHORE.
FT LIPID 314 314 PALMITATE (BY SIMILARITY).
FT LIPID 315 315 PALMITATE (BY SIMILARITY).
FT CONFLICT 28 28 Q -> W (IN REF. 3).
FT CONFLICT 135 135 K -> R (IN REF. 2).
FT CONFLICT 163 163 A -> T (IN REF. 2).
SQ SEQUENCE 351 AA; 38173 MW; 56BDFAD187008338 CRC64;

Query Match 72.0%; Score 36; DB 1; Length 351;
Best Local Similarity 75.0%; Pred. No. 21;
Matches 6; Conservative 1; Mismatches 1; Indels 1; Gaps 0;

QY 1 STQTPENG 8
DB 5 SSQAPPNG 12

RESULT 2
RF1_AERPE STANDARD; PRT; 373 AA.
AC QYAF1;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Peptide chain release factor subunit 1 (translation termination factor arf1).
DE ARF1.
GN PRF1 OR APE1988.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE=99310339; PubMed=10392966;
RA Kwarabavasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic crenarchaeon, Aeropyrum pernix K1.";
```

```

RL DNA Res. 6:83-101(1999).
CC -1- FUNCTION: Directs the termination of nascent peptide synthesis
CC (translation) in response to the termination codons UAA, UAG and
CC UGA (By similarity).
CC -1- SUBUNIT: Heterodimer of two subunits, one of which binds GTP (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -1- SIMILARITY: BELONGS TO THE EUKARYOTIC RELEASE FACTOR 1 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; AP000063; BAA80998.1; ALT_INIT.
CC HSSP; P46055; 1DT9.
CC HAMAP; MF_00424; 1.
CC InterPro; IPR005140; eRF1_1.
CC InterPro; IPR005141; eRF1_2.
CC InterPro; IPR005142; eRF1_3.
CC Pfam; PF03463; eRF1_1; 1.
CC Pfam; PF03464; eRF1_2; 1.
CC Pfam; PF03465; eRF1_3; 1.
KW Protein biosynthesis; Complete proteome.
SQ SEQUENCE 373 AA; 41950 MW; 6BFD39299E876477 CRC64;

Query Match 72.0%; Score 36; DB 1; Length 373;
Best Local Similarity 85.7%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TQTPENG 8
DB 86 TSTPPNG 92

RESULT 3
YX70_CORGL STANDARD; PRT; 718 AA.
AC P54122;
DT 01-OCT-1996 (Rel. 34, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein Cgl1970.
GN CGL1970.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Patek M., Bilic M., Krumbach K., Eikmanns B., Sahm H., Eggeling L.;
RT "Identification and transcriptional analysis of the dapB-CRF2-dapA-
RT ORF4 operon of Corynebacterium glutamicum, encoding two enzymes
RT involved in L-lysine synthesis.";
RL Biotechnol. Lett. 19:1113-1117(1997).
RN [2]
RP REVISIONS.
RA Eggeling L.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE UPF0036 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
```


CC serpin does not inhibit activated protein C (APC), does not require
 CC activation by heparin, and includes a target sequence which interacts
 CC with the proteolytically active site of thrombin to inhibit its
 CC activity. The inhibitor protein can be used as antithrombotic, or
 CC anticoagulant agents for treating or preventing disseminated
 CC intravascular coagulation (DIC), unstable angina, myocardial infarction,
 CC thrombotic stroke, thrombosis, pulmonary embolism, or other clotting
 CC disorders. When this protein is labelled, it can be used in diagnostic
 CC assays to determine the presence of serine proteases, especially
 CC thrombin. The protein has the same specificity as AT-III, but does not
 CC need to be administered with heparin, and is approximately 1500 times
 CC more active.
 XX

SQ Sequence 19 AA;

Query Match 70.5%; Score 31; DB 17; Length 19;
 Best Local Similarity 77.8%; Pred. No. 39;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 STAVFFAGR 9
 |||||
 Db 4 STAVVIAGR 12

RESULT 10

AAP80061
 ID AAP80061 standard; protein; 38 AA.

XX AAP80061;

XX 05-NOV-1990 (first entry)

XX Sequence of reactive centre of antithrombin-III.

XX Protease inhibitor; thrombolytic.

XX Homo sapiens.

XX DE3713272-A.

XX 03-NOV-1988.

XX 18-APR-1987; 87DE-3713272.

XX 18-APR-1987; 87DE-3713272.

XX (BEHW) BEHRINGER AG.

XX Lecander I, Astedt B, Ny T;

XX WPI; 1988-315694/45.

XX New DNA sequence encoding plasminogen activator inhibitor type 2 -
 XX PT and derived proteins and antibodies, useful therapeutically and
 XX PT in diagnosis

XX Disclosure; Table 2 Page 6; 6pp; German.

XX The table compares the AA SQs from the reactive centres of PAII
 CC (AAP80059), alpha-1 antitrypsin (AAP80060), antithrombin-III (AAP80061)
 CC and PA12 (AAP80062) PA12 is a strong and specific inhibitor for the
 CC active forms of PA so is used to treat or prevent disorders associated
 CC with abnormal levels of PA. PA is implicated in haemostasis, tissue
 CC repair, metastasis formation, ovulation, fertility, macrophage migration
 CC and inflammatory processes.

XX Sequence 38 AA;

Query Match 70.5%; Score 31; DB 9; Length 38;

Best Local Similarity 77.8%; Pred. No. 76;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 STAVFFAGR 9

Db 25 STAVVIAGR 33
 |||||
 |||||

RESULT 11

AAP90537
 ID AAP90537 standard; peptide; 48 AA.

XX AAP90537;

XX 25-MAR-2003 (updated)

DT 31-OCT-2002 (updated)

DT 19-OCT-1989 (first entry)

XX Reactive site of antithrombin-III.

XX Antithrombin-III; reactive site; alpha-2-plasmin inhibitor.

XX Synthetic.

XX EP326013-A.

XX 02-AUG-1989.

XX 19-JAN-1989; 89EP-0100851.

XX 28-JAN-1988; 88JP-0015762.

XX (FARH) HOECHST AG.

XX Toba M, Tone M, Kikuno R, Hashimoto T;

XX WPI; 1989-221985/31.

XX N-PSDB; AAN90369.

XX Alpha-2 plasmin inhibitor deriv. - having reactive site cleavable with
 XX PT plasmin deleted or replaced to give protein active in inhibiting blood
 XX PT coagulation.

XX Disclosure; Page -; 16pp; English.

XX The peptide is the reactive site of antithrombin-III (residues 376-423),
 CC an antithrombin-active serine protease inhibitor. It may be used to
 CC replace the active site of alpha-2-plasmin inhibitor (see AAN90369 and
 CC AAP90534).

XX (Updated on 31-OCT-2002 to add missing OS field.)

XX (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 48 AA;

Query Match 70.5%; Score 31; DB 10; Length 48;

Best Local Similarity 77.8%; Pred. No. 95;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 STAVFFAGR 9
 |||||
 |||||

Db 10 STAVVIAGR 18

RESULT 12

AAM92770
 ID AAM92770 standard; Protein; 71 AA.

XX AAM92770;

XX 06-NOV-2001 (first entry)

XX Human digestive system antigen SEQ ID NO: 2119.

XX Human; digestive system antigen; Gene therapy; cancer; appendicitis;
 XX ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
 XX digestive system disorder; Meckel's diverticulum.

Qy

XX Rosen CA, Barash SC, Ruben SM;
 XX WPI; 2001-502630/55.
 DR N-PSDB; AAK89543.
 XX Polynucleotides encoding digestive system antigens, useful for
 PT diagnosing, treating, preventing and/or prognosing disorders of the
 PT digestive system, particularly cancer and cancer metastases -
 XX
 PS Claim 11; SEQ ID NO 2119; 986pp; English.
 XX
 CC The present invention provides the protein and coding sequences of a
 CC number of human digestive system antigens. These can be used in the
 CC diagnosis, treatment and prevention of digestive system disorders,
 CC including cancer, Meckel's diverticulum, bacterial or parasitic
 CC infections, appendicitis, Hirschsprung's disease, chronic colitis or
 CC ulcerative colitis. The present sequence is a digestive system antigen of
 CC the invention.
 XX
 SQ Sequence 71 AA;
 Query Match 70.5%; Score 31; DB 22; Length 71;
 Best Local Similarity 85.7%; Pred. NO. 1.4e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 TAVFFAG 8
 |||||
 DB 5 TAVFFCG 11
 RESULT 13
 AAW38477
 ID AAW38477 standard; Protein; 110 AA.
 AC AAW38477;
 XX
 DT 06-NOV-1998 (first entry)
 DE
 XX Streptococcus pneumoniae protein of unknown function.
 XX
 KW Streptococcus pneumoniae protein; genetic immunisation; antagonist;
 KW immunological response; inoculation; antibody production; inhibitor;
 KW T cell immune response; antimicrobial compound; bacterial adhesion;
 KW extracellular matrix protein; protein-mediated cell invasion; wound;
 KW pathogenesis.
 XX
 OS Streptococcus pneumoniae.
 XX
 FN WO9743303-A1.
 XX
 PD 20-NOV-1997.
 XX
 PF 14-MAY-1997; 97WO-US07950.
 XX
 PR 14-MAY-1996; 96US-0017670.
 XX
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 PA (SMIK) SMITHKLINE BEECHAM PLC.
 XX
 PI Black MT, Hodgson JE, Knowles DJC, Nicholas RO;
 PI Stodola RK;
 XX
 DR WPI; 1998-008793/01.
 DR N-PSDB; AAT98545.
 XX
 PS Novel Streptococcus pneumoniae proteins and related DNA - useful for
 PT diagnosing anti-microbial agents for treatment of bacterial
 PT infections
 XX
 PS Claim 12; Page 270; 483pp; English.
 XX
 CC This sequence represents a Streptococcus pneumoniae protein of

CC unknown function, and is encoded by a DNA sequence of the invention.
 CC The DNA sequences were isolated from Streptococcus pneumoniae strain
 CC 0100993 (NCIMB 40794). The Streptococcus pneumoniae proteins of the
 CC invention can be used to identify compounds which interact with and
 CC inhibit or activate the activity of the proteins. Antagonists can be
 CC used to treat diseases caused by S. pneumoniae proteins, through genetic
 CC immunisation. They can also be used to induce an immunological response
 CC in a mammal by inoculation with the S. pneumoniae proteins or delivery
 CC of the encoding nucleic acids in a vector adequate to produce antibody
 CC and/or T cell immune responses to protect the animal from disease. The
 CC proteins can also be used to identify antimicrobial compounds which are
 CC capable of inhibiting their bioactivity. In particular the proteins of
 CC the invention can be used to prevent adhesion of bacteria to mammalian
 CC extracellular matrix proteins on in-dwelling devices or in wounds, to
 CC block protein-mediated mammalian cell invasion, and to block the normal
 CC progression of pathogenesis in infections initiated other than by the
 CC implantation of in-dwelling devices or other surgical techniques.
 XX
 SQ Sequence 110 AA;
 Query Match 70.5%; Score 31; DB 19; Length 110;
 Best Local Similarity 75.0%; Pred. NO. 2.1e+02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 2 TAVFFAG 9
 |||||
 DB 28 TATFFIGR 35
 RESULT 14
 ABU01952
 ID ABU01952 standard; Protein; 115 AA.
 AC ABU01952;
 XX
 DT 11-FEB-2003 (first entry)
 DE
 XX S. pneumoniae type 4 strain protein from coding region #1529.
 XX
 KW Bacterial meningitis; pneumonia; sepsis; otitis media;
 KW ear infection; antinflammatory; antibacterial; immunostimulant;
 KW auditory; respiratory; gene therapy; vaccine.
 XX
 OS Streptococcus pneumoniae type 4 strain.
 XX
 FN WO200277021-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 27-MAR-2002; 2002WO-IB02163.
 XX
 PR 27-MAR-2001; 2001GB-0007658.
 XX
 PA (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Masignani V, Tettelin H, Fraser C;
 XX
 DR WPI; 2003-040579/03.
 DR N-PSDB; ABX07241.
 XX
 PT New proteins and nucleic acid molecules from Streptococcus pneumoniae,
 PT useful as medicaments for treating or preventing a disease or infection
 PT due to streptococcus bacteria, such as pneumonia, sepsis, otitis media
 PT or ear infection -
 XX
 PS Claim 1; SEQ ID No 3058; 56pp; English.
 XX
 CC The invention relates to a protein comprising or having at least 50%
 CC identity to any of the 2469 amino acid sequences, identified in the
 CC specification (available on a computer readable format), or its fragment,
 CC expressed from 2469 of 2489 identified DNA coding regions from the
 CC Streptococcus pneumoniae type 4 strain genomic sequence appearing as


```
Query Match 72.7%; Score 32; DB 4; Length 35;
Best Local Similarity 62.5%; Pred. No. 9.3;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 TAVFFAGR 9
DB 3 TAVFCGR 10

RESULT 3
US-09-271-608-13
; Sequence 13, Application US/09271608C
; Patent No. 6245741
; GENERAL INFORMATION:
; APPLICANT: Broze, George J., Jr.
; TITLE OF INVENTION: Protein Z-Dependent Protease Inhibitor
; FILE REFERENCE: WU-3110/1
; CURRENT APPLICATION NUMBER: US/09/271,608C
; CURRENT FILING DATE: 1999-03-17
; EARLIER APPLICATION NUMBER: US 60/086,571
; EARLIER FILING DATE: 1998-05-19
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Word Perfect 5.0
; SEQ ID NO 13
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: /note="synthetic construct"
US-09-271-608-13

Query Match 70.5%; Score 31; DB 3; Length 56;
Best Local Similarity 77.8%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 2; Indels 2; Gaps 0;

QY 1 STAVFFAGR 9
DB 9 STAVVIAGR 17

RESULT 4
US-09-695-950-13
; Sequence 13, Application US/09695950
; Patent No. 6265378
; GENERAL INFORMATION:
; APPLICANT: Broze, George J., Jr.
; TITLE OF INVENTION: Protein Z-Dependent Protease Inhibitor
; FILE REFERENCE: WU-3110/1
; CURRENT APPLICATION NUMBER: US/09/695,950
; CURRENT FILING DATE: 2000-10-25
; PRIOR APPLICATION NUMBER: 09/271,608
; PRIOR FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Word Perfect 5.0
; SEQ ID NO 13
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: /note="synthetic construct"
US-09-695-950-13

Query Match 70.5%; Score 31; DB 3; Length 56;
Best Local Similarity 77.8%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 2; Indels 2; Gaps 0;

QY 1 STAVFFAGR 9
DB 9 STAVVIAGR 17

RESULT 5
US-09-696-147-13
; Sequence 13, Application US/09696147
; Patent No. 6271367
; GENERAL INFORMATION:
; APPLICANT: Broze, George J., Jr.
; TITLE OF INVENTION: Protein Z-Dependent Protease Inhibitor
; FILE REFERENCE: WU-3110/1
; CURRENT APPLICATION NUMBER: US/09/696,147
; CURRENT FILING DATE: 2000-10-25
; PRIOR APPLICATION NUMBER: 09/271,608
; PRIOR FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Word Perfect 5.0
; SEQ ID NO 13
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: /note="synthetic construct"
US-09-696-147-13

Query Match 70.5%; Score 31; DB 3; Length 56;
Best Local Similarity 77.8%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 2; Indels 2; Gaps 0;

QY 1 STAVFFAGR 9
DB 9 STAVVIAGR 17

RESULT 6
US-09-696-364-13
; Sequence 13, Application US/09696364
; Patent No. 6369031
; GENERAL INFORMATION:
; APPLICANT: Broze, George J., Jr.
; TITLE OF INVENTION: Protein Z-Dependent Protease Inhibitor
; FILE REFERENCE: WU-3110/1
; CURRENT APPLICATION NUMBER: US/09/696,364
; CURRENT FILING DATE: 2000-10-25
; PRIOR APPLICATION NUMBER: 09/271,608
; PRIOR FILING DATE: 1999-03-17
; PRIOR APPLICATION NUMBER: US 60/086,571
; PRIOR FILING DATE: 1998-05-19
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Word Perfect 5.0
; SEQ ID NO 13
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: /note="synthetic construct"
US-09-696-364-13

Query Match 70.5%; Score 31; DB 4; Length 56;
Best Local Similarity 77.8%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 2; Indels 2; Gaps 0;

QY 1 STAVFFAGR 9
DB 9 STAVVIAGR 17

RESULT 7
US-08-858-207A-273
; Sequence 273, Application US/08858207A
; Patent No. 6348328
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Nicholas, Richard
; APPLICANT: Stodola, Robert
; TITLE OF INVENTION: No. 6348328el Compounds
```

```
Query Match          70.5%; Score 31; DB 1; Length 464;
Best Local Similarity 77.8%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 STAVFFAGR 9
DB      417 STAVVIAGR 425

RESULT 10
US-09-556-877-192
; Sequence 192, Application US/09556877
; Patent No. 6432916
; GENERAL INFORMATION:
; APPLICANT: Probst, Peter
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir
; APPLICANT: Fling, Steve
; APPLICANT: Maisonneuve, Jeff
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; FILE REFERENCE: 210121.469C5
; CURRENT APPLICATION NUMBER: US/09/556,877
; CURRENT FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 192
; LENGTH: 848
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-556-877-192

Query Match          70.5%; Score 31; DB 4; Length 848;
Best Local Similarity 55.6%; Pred. No. 3.3e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1 STAVFFAGR 9
DB      243 SDAIFYAGK 251

RESULT 11
US-09-620-412C-192
; Sequence 192, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
; APPLICANT: Steven P. Fling
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; FILE REFERENCE: 210121.469C7
; CURRENT APPLICATION NUMBER: US/09/620,412C
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 192
; LENGTH: 848
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-620-412C-192

Query Match          70.5%; Score 31; DB 4; Length 848;
Best Local Similarity 55.6%; Pred. No. 3.3e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1 STAVFFAGR 9
DB      243 SDAIFYAGK 251

RESULT 12
US-09-598-419-192
; Sequence 192, Application US/09598419
```

```
Patent No. 6565856
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Scholler, John
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
; FILE REFERENCE: 210121.469C6
; CURRENT APPLICATION NUMBER: US/09/598,419
; CURRENT FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 357
; SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 192
; LENGTH: 848
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-598-419-192

Query Match          70.5%; Score 31; DB 4; Length 848;
Best Local Similarity 55.6%; Pred. No. 3.3e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1 STAVFFAGR 9
DB      243 SDAIFYAGK 251

RESULT 13
US-09-556-877-178
; Sequence 178, Application US/09556877
; Patent No. 6432916
; GENERAL INFORMATION:
; APPLICANT: Probst, Peter
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir
; APPLICANT: Fling, Steve
; APPLICANT: Maisonneuve, Jeff
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; FILE REFERENCE: 210121.469C5
; CURRENT APPLICATION NUMBER: US/09/556,877
; CURRENT FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 178
; LENGTH: 1530
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-556-877-178

Query Match          70.5%; Score 31; DB 4; Length 1530;
Best Local Similarity 55.6%; Pred. No. 5.9e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1 STAVFFAGR 9
DB      925 SDAIFYAGK 933

RESULT 14
US-09-620-412C-178
; Sequence 178, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
; APPLICANT: Steven P. Fling
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; FILE REFERENCE: 210121.469C7
; CURRENT APPLICATION NUMBER: US/09/620,412C
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 178
; LENGTH: 1530
```

Publication No. US20030143224A1
 GENERAL INFORMATION:
 APPLICANT: Pusiner, Stanley
 APPLICANT: Safar, Jiri
 APPLICANT: Williamson, Anthony
 APPLICANT: Burton, Dennis
 TITLE OF INVENTION: Antibodies Specific for Ungulate PrP
 FILE REFERENCE: UCAL-194
 CURRENT APPLICATION NUMBER: US/10/355,780
 PRIOR FILING DATE: 2003-01-30
 PRIOR APPLICATION NUMBER: US/09/627,218B
 PRIOR FILING DATE: 2000-07-27
 NUMBER OF SEQ ID NOS: 11
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 7
 LENGTH: 35
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: synthesized peptide
 US-10-355-780-7

Query Match 72.7%; Score 32; DB 12; Length 35;
 Best Local Similarity 62.5%; Pred. No. 13;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TAVFFAGR 9
 Db 3 TATYFCGR 10

RESULT 3
 US-10-156-761-9311
 Sequence 9311, Application US/10156761
 Publication No. US20030119018A1
 GENERAL INFORMATION:
 APPLICANT: OMURA, SATOSHI
 APPLICANT: IKEDA, HARUO
 APPLICANT: ISHIKAWA, JUN
 APPLICANT: HORIKAWA, HIROSHI
 APPLICANT: SHIBA, TADAYOSHI
 APPLICANT: SAKAKI, YOSHIYUKI
 APPLICANT: HATTORI, MASAHIRA
 TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 FILE REFERENCE: 249-262
 CURRENT APPLICATION NUMBER: US/10/156,761
 PRIOR FILING DATE: 2002-05-29
 PRIOR APPLICATION NUMBER: JP 2001-204089
 PRIOR FILING DATE: 2001-05-30
 PRIOR APPLICATION NUMBER: JP 2001-272697
 PRIOR FILING DATE: 2001-08-02
 NUMBER OF SEQ ID NOS: 15109
 SEQ ID NO 9311
 LENGTH: 203
 TYPE: PRT
 ORGANISM: Streptomyces avermitilis
 US-10-156-761-9311

Query Match 72.7%; Score 32; DB 15; Length 203;
 Best Local Similarity 77.8%; Pred. No. 78;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 STAVFFAGR 9
 Db 38 SLAPFFAGR 46

RESULT 4
 US-10-043-452-13
 Sequence 13, Application US/10043452
 Publication No. US2002018254A1
 GENERAL INFORMATION:
 APPLICANT: Broze, George J., Jr.

TITLE OF INVENTION: Protein Z-Dependent Protease Inhibitor
 FILE REFERENCE: WU-3110/1
 CURRENT APPLICATION NUMBER: US/10/043,452
 CURRENT FILING DATE: 2002-01-10
 PRIOR APPLICATION NUMBER: 09/696,364
 PRIOR FILING DATE: 2000-10-25
 PRIOR APPLICATION NUMBER: 09/271,608
 PRIOR FILING DATE: 1999-03-17
 PRIOR APPLICATION NUMBER: US 60/086,571
 PRIOR FILING DATE: 1998-05-19
 NUMBER OF SEQ ID NOS: 16
 SOFTWARE: Word Perfect 5.0
 SEQ ID NO 13
 LENGTH: 56
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: /note="synthetic construct"
 US-10-043-452-13

Query Match 70.5%; Score 31; DB 14; Length 56;
 Best Local Similarity 77.8%; Pred. No. 33;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 STAVFFAGR 9
 Db 9 STAVIAGR 17

RESULT 5
 US-10-156-761-12838
 Sequence 12838, Application US/10156761
 Publication No. US20030119018A1
 GENERAL INFORMATION:
 APPLICANT: OMURA, SATOSHI
 APPLICANT: IKEDA, HARUO
 APPLICANT: ISHIKAWA, JUN
 APPLICANT: HORIKAWA, HIROSHI
 APPLICANT: SHIBA, TADAYOSHI
 APPLICANT: SAKAKI, YOSHIYUKI
 APPLICANT: HATTORI, MASAHIRA
 TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 FILE REFERENCE: 249-262
 CURRENT APPLICATION NUMBER: US/10/156,761
 CURRENT FILING DATE: 2002-05-29
 PRIOR APPLICATION NUMBER: JP 2001-204089
 PRIOR FILING DATE: 2001-05-30
 PRIOR APPLICATION NUMBER: JP 2001-272697
 PRIOR FILING DATE: 2001-08-02
 NUMBER OF SEQ ID NOS: 15109
 SEQ ID NO 12838
 LENGTH: 306
 TYPE: PRT
 ORGANISM: Streptomyces avermitilis
 US-10-156-761-12838

Query Match 70.5%; Score 31; DB 15; Length 306;
 Best Local Similarity 85.7%; Pred. No. 1.9e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 AVFFAGR 9
 Db 47 AVFFVGR 53

RESULT 6
 US-09-414-834-1
 Sequence 1, Application US/09414834
 Patent No. US2002076413A1
 GENERAL INFORMATION:
 APPLICANT: O'Reilly, Michael S.
 APPLICANT: Polkman, M. Judah
 APPLICANT: Pirie-Shepherd, Steven

```
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, KIROSHI
; APPLICANT: SHIERA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 10429
; LENGTH: 3970
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-10429

Query Match 70.5%; Score 31; DB 15; Length 3970;
Best Local Similarity 85.7%; Pred. No. 2.7e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 AVFFAG 9
DB 924 AAFVAG 930

RESULT 12
US-10-115-223-19
; Sequence 19, Application US/10115223
; Publication No. US20030176334A1
; GENERAL INFORMATION:
; APPLICANT: Brooks, Peter
; APPLICANT: Cheresch, David A.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS USEFUL FOR INHIBITION OF
; TITLE OF INVENTION: ANGIOGENESIS
; FILE REFERENCE: MER00495
; CURRENT APPLICATION NUMBER: US/10/115,223
; CURRENT FILING DATE: 2002-04-02
; PRIOR APPLICATION NUMBER: US/09/194,468
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: 60/018,773
; PRIOR FILING DATE: 1996-05-31
; PRIOR APPLICATION NUMBER: 60/015,896
; PRIOR FILING DATE: 1996-05-31
; PRIOR APPLICATION NUMBER: PCT/US97/09158
; PRIOR FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 19
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-115-223-19

Query Match 68.2%; Score 30; DB 12; Length 74;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AVFFAG 8
DB 65 AVFFAG 70

RESULT 13
US-10-115-223-25
; Sequence 25, Application US/10115223
; Publication No. US20030176334A1
; GENERAL INFORMATION:
; APPLICANT: Brooks, Peter
; APPLICANT: Cheresch, David A.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS USEFUL FOR INHIBITION OF
; TITLE OF INVENTION: ANGIOGENESIS
; FILE REFERENCE: MER00495
; CURRENT APPLICATION NUMBER: US/10/115,223
; CURRENT FILING DATE: 2002-04-02
; PRIOR APPLICATION NUMBER: US/09/194,468
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: 60/018,773
; PRIOR FILING DATE: 1996-05-31
; PRIOR APPLICATION NUMBER: 60/015,896
; PRIOR FILING DATE: 1996-05-31
; PRIOR APPLICATION NUMBER: PCT/US97/09158
; PRIOR FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 19
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-115-223-19

Query Match 68.2%; Score 30; DB 12; Length 74;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AVFFAG 8
DB 65 AVFFAG 70

RESULT 14
US-10-115-223-20
; Sequence 20, Application US/10115223
; Publication No. US20030176334A1
; GENERAL INFORMATION:
; APPLICANT: Brooks, Peter
; APPLICANT: Cheresch, David A.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS USEFUL FOR INHIBITION OF
; TITLE OF INVENTION: ANGIOGENESIS
; FILE REFERENCE: MER00495
; CURRENT APPLICATION NUMBER: US/10/115,223
; CURRENT FILING DATE: 2002-04-02
; PRIOR APPLICATION NUMBER: US/09/194,468
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: 60/018,773
; PRIOR FILING DATE: 1996-05-31
; PRIOR APPLICATION NUMBER: 60/015,896
; PRIOR FILING DATE: 1996-05-31
; PRIOR APPLICATION NUMBER: PCT/US97/09158
; PRIOR FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 20
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-115-223-20

Query Match 68.2%; Score 30; DB 12; Length 108;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AVFFAG 8
DB 65 AVFFAG 70

RESULT 15
US-10-115-223-26
; Sequence 26, Application US/10115223
; Publication No. US20030176334A1
; GENERAL INFORMATION:
; APPLICANT: Brooks, Peter
; APPLICANT: Cheresch, David A.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS USEFUL FOR INHIBITION OF
; TITLE OF INVENTION: ANGIOGENESIS
; FILE REFERENCE: MER00495
; CURRENT APPLICATION NUMBER: US/10/115,223
; CURRENT FILING DATE: 2002-04-02
; PRIOR APPLICATION NUMBER: US/09/194,468
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: 60/018,773
; PRIOR FILING DATE: 1996-05-31
; PRIOR APPLICATION NUMBER: 60/015,896
; PRIOR FILING DATE: 1996-05-31
; PRIOR APPLICATION NUMBER: PCT/US97/09158
; PRIOR FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 25
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Gallus gallus
US-10-115-223-25

Query Match 68.2%; Score 30; DB 12; Length 74;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AVFFAG 8
DB 65 AVFFAG 70
```

```
Best Local Similarity 66.7%; Pred. No. 9.6;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 STAVFFAGR 9
   ||:|:|
Db 91 STSMFYAGR 99

RESULT 3
E72415
Zinc ABC transporter, permease protein - Thermotoga maritima (strain MSB8)
C/Species: Thermotoga maritima
C/Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C/Accession: E72415
R/Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
Nature 399, 323-329, 1999
A/Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A/Reference number: A72200; MUID:99287316; PMID:10360571
A/Accession: E72415
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-282 <ARN>
A/Cross-references: GB:AE001698; GB:AE000512; NID:G4980609; PIDN:AAD35219.1; PID:G498061
A/Experimental source: strain MSB8
C/Genetics:
A/Gene: TW0125
C/Superfamily: conserved hypothetical protein H10360

Query Match 79.5%; Score 35; DB 2; Length 282;
Best Local Similarity 87.5%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 STAVFFAGR 8
   ||:|:|
Db 245 STGVFFAG 252

RESULT 4
S58121
19 heavy chain V region precursor - human (fragment)
C/Species: Homo sapiens (man)
C/Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 21-Jan-2000
C/Accession: S58121
R/Kueppers, R.
submitted to the EMBL Data Library, July 1995
A/Reference number: S58121
A/Accession: S58121
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-126 <KUE>
A/Cross-references: EMBL:X89595; NID:G929650; PIDN:CAA61756.1; PID:G929651
C/Genetics:
A/Introns: 9/1
C/Superfamily: immunoglobulin V region; immunoglobulin homology
C/Keywords: heterotrimer; immunoglobulin
F;27-110/Domain: immunoglobulin homology <IMM>

Query Match 75.0%; Score 33; DB 2; Length 126;
Best Local Similarity 75.0%; Pred. No. 15;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 TAVFFAGR 9
   ||:|:|
Db 103 TAVYFCGR 110

RESULT 5
S11920
PulO protein - Klebsiella oxytoca
C/Species: Klebsiella oxytoca
C/Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 11-Jan-2000
```

```
C/Accession: S11920
R/Reyss, I.; Pugsley, A.P.
Mol. Gen. Genet. 222, 176-184, 1990
A/Title: Five additional genes in the pulC-O operon of the gram-negative bacterium Kleb:
A/Reference number: S11917; MUID:91109698; PMID:2129543
A/Accession: S11920
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-198 <REY>
C/Superfamily: secretion protein xcpW

Query Match 75.0%; Score 33; DB 2; Length 198;
Best Local Similarity 66.7%; Pred. No. 23;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 STAVFFAGR 9
   ||:|:|
Db 73 SASLFFAGR 81

RESULT 6
H75521
cytochrome c-type biogenesis protein, heme exporter protein B - Deinococcus radiodurans
C/Species: Deinococcus radiodurans
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C/Accession: H75521
R/White, O.; Bisen, J.A.; Heidberg, J.P.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
S.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; M
Science 286, 1571-1577, 1999
A/Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A/Reference number: A75250; MUID:20036896; PMID:10567266
A/Accession: H75521
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-221 <WHI>
A/Cross-references: GB:AE001901; GB:AE000513; NID:G6458087; PIDN:AAF09987.1; PID:G64580
A/Experimental source: strain R1
C/Genetics:
A/Gene: DR0407
C/Map position: 1
C/Superfamily: cytochrome c biogenesis protein CycW

Query Match 72.7%; Score 32; DB 2; Length 221;
Best Local Similarity 75.0%; Pred. No. 40;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 STAVFFAG 8
   ||:|:|
Db 23 STTLFFAG 30

RESULT 7
E70391
major facilitator family transporter - Aquifex aeolicus
C/Species: Aquifex aeolicus
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
C/Accession: E70391
R/Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O
V.
Nature 392, 353-358, 1998
A/Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A/Reference number: A70300; MUID:98196666; PMID:9537320
A/Accession: E70391
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-511 <AQF>
A/Cross-references: GB:AE000721; GB:AE000657; NID:G2983544; PIDN:AA07127.1; PID:G29835
A/Experimental source: strain VF5
C/Genetics:
A/Gene: emrB
C/Superfamily: lincomycin-resistance protein lmrB
```

A;Residues: 1-115 <KUR>
A;Cross-references: GB:AE007317; PIDN:AAL00156.1; PID:g15458999; GSPDB:GN00174
C;Genetics:
A;Gene: bta

Query Match 70.5%; Score 31; DB 2; Length 115;
Best Local Similarity 75.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TAVFFAG 9
||| |||
DB 28 TATFFGR 35

RESULT 13
A69946
hypothetical protein yqar - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
C;Accession: A69946
R;Kunst, P.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertoni, R.; Bron, S.; Brouillette, S.; Bruch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chao, A.; Ehrlich, S.D.; Emmerman, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.; Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritze, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallet, A.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Iech, J.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, Koetter, P.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, Y.; M.; Ogiwara, K.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon, A;Authors: Schleicher, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron, akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; PMID:98044033; PMID:9384377
A;Accession: A69946
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-154 <KUN>
A;Cross-references: GB:Z99117; GB:AL009126; NID:g2634966; PIDN:CAB14562.1; PID:g2635066
A;Experimental source: strain 168
C;Genetics:
A;Gene: yqar
C;Superfamily: Bacillus subtilis hypothetical protein yqar

Query Match 70.5%; Score 31; DB 2; Length 154;
Best Local Similarity 85.7%; Pred. No. 46;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TAVFFAG 8
||| |||
DB 15 TAVFFTG 21

RESULT 14
T01205
sulfate transport protein - maize (fragment)
N;Alternate names: sulfate permease
C;Species: Zea mays (maize)
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 26-May-2000
C;Accession: T01205
R;Bolchi, A.; Petrucci, S.; Foroni, C.; Tenca, G.L.; Ottonello, S.
submitted to the EMBL Data Library, July 1997
A;Description: Sulfate permease and ATP sulfurylase mRNA are coordinately modulated in m
A;Reference number: 214260
A;Accession: T01205
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-233 <BOI>
A;Cross-references: EMBL:AF016306; NID:g2738751; PIDN:AAB94543.1; PID:g2738752
A;Experimental source: cultivar Paulo; root
C;Superfamily: sulfate transport protein

C;Keywords: sulfate transport

Query Match 70.5%; Score 31; DB 2; Length 233;
Best Local Similarity 85.7%; Pred. No. 67;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TAVFFAG 8
||| |||
DB 35 TATFFAG 41

RESULT 15

AI2709
branched-chain amino acid permease Atu1082 [imported] - Agrobacterium tumefaciens (stra.
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C;Accession: AI2709
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.F.; Chen, Y.; Woo,
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, S.W.

A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; PMID:21608550; PMID:11743193
A;Accession: AI2709
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-234 <KUR>
A;Cross-references: GB:AE008688; PIDN:AAL42095.1; PID:g17739477; GSPDB:GN00186
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atu1082
A;Map position: circular chromosome
C;Superfamily: hypothetical protein b2682

Query Match 70.5%; Score 31; DB 2; Length 234;
Best Local Similarity 85.7%; Pred. No. 67;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TAVFFAG 8
||| |||
DB 129 TAAFFAG 135

Search completed: December 11, 2003, 18:29:47
Job time : 11.3333 secs

Db 91 STSMFYAG 99

RESULT 2

Y125 THMA STANDARD; PRT; 282 AA.

AC Q9MX9; ID Y125 THMA STANDARD; PRT; 282 AA.

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Probable metal transport system membrane protein TM0125.

GN TM0125.

OS Thermotoga maritima.

OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.

OX NCBI_TaxID=2336; [1]_TaxID=2336;

RN SEQUENCE FROM N.A.

RP STRAIN=MSB8 / DSM 3109;

RC MEDLINE=9287316; PubMed=10360571;

RX Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A., McDonald L., Utterback T.R., Malek J.A., Lither K.D., Garrett M.M., Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D., Reidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O., Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M., and Bacteria from RT "Evidence for lateral gene transfer between Archaea and Bacteria from RL genome sequence of Thermotoga maritima."

CC -1- FUNCTION: PART OF AN ATP-DRIVEN TRANSPORT SYSTEM

CC Nature 399:323-329 (1999).

CC -1- TM0123/TM0124/TM0125 FOR A METAL.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane (Probable).

CC -1- SIMILARITY: BELONGS TO THE ABC-3 SUBFAMILY OF INTEGRAL MEMBRANE PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; M32613; AAA25132.1; --

DR InterPro: IPR001120; Prok N methyltn.

DR PROSITE; PS00409; PROKAR_NTER_METHYL; 1.

KW Transport; Methylation.

FT PROPEP 1 7 BY SIMILARITY.

FT CHAIN 8 198 GENERAL SECRETION PATHWAY PROTEIN J.

FT MOD RES 8 8 METHYLATION (BY SIMILARITY).

SQ SEQUENCE 198 AA; 22206 MW; 956977D5EF4E410D CRC64;

Query Match 75.0%; Score 33; DB 1; Length 198;

Best Local Similarity 66.7%; Pred. No. 8.3;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 STAVFFAGR 9

Db 73 SASLFFAGR 81

RESULT 4

YQAR_BACSU ID YQAR_BACSU STANDARD; PRT; 154 AA.

AC P45914;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Hypothetical protein yqar.

GN YQAR.

OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OX NCBI_TaxID=1423; [1]_TaxID=1423;

RN SEQUENCE FROM N.A.

RP STRAIN=168 / JH642;

RC MEDLINE=95219086; PubMed=7704261;

RX Takemaru K.-I., Mizuno M., Sato T., Takeuchi M., Kobayashi Y.; RT "Complete nucleotide sequence of a skin element excised by DNA RT rearrangement during sporulation in Bacillus subtilis.";

RL Microbiology 141:323-327(1995).

RN [2]

RN SEQUENCE FROM N.A.

RC STRAIN=168 / JH642;

Db 91 STSMFYAG 99

RESULT 3

GSPJ_KLEPN ID GSPJ_KLEPN STANDARD; PRT; 198 AA.

AC Q9MX9; ID Y125 THMA STANDARD; PRT; 282 AA.

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Probable metal transport system membrane protein TM0125.

GN TM0125.

OS Thermotoga maritima.

OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.

OX NCBI_TaxID=2336; [1]_TaxID=2336;

RN SEQUENCE FROM N.A.

RP STRAIN=MSB8 / DSM 3109;

RC MEDLINE=9287316; PubMed=10360571;

RX Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A., McDonald L., Utterback T.R., Malek J.A., Lither K.D., Garrett M.M., Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D., Reidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O., Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M., and Bacteria from RT "Evidence for lateral gene transfer between Archaea and Bacteria from RL genome sequence of Thermotoga maritima."

CC -1- FUNCTION: PART OF AN ATP-DRIVEN TRANSPORT SYSTEM

CC Nature 399:323-329 (1999).

CC -1- TM0123/TM0124/TM0125 FOR A METAL.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane (Probable).

CC -1- SIMILARITY: BELONGS TO THE ABC-3 SUBFAMILY OF INTEGRAL MEMBRANE PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; AE001698; AAD35219.1; --

DR InterPro: IPR001626; ABC_transp3.

DR Pfam; PF00950; ABC-3; 1.

KW Hypothetical protein; Transport; Transmembrane; Inner membrane; Complete proteome.

FT TRANSMEM 33 53 POTENTIAL.

FT TRANSMEM 58 78 POTENTIAL.

FT TRANSMEM 79 99 POTENTIAL.

FT TRANSMEM 109 129 POTENTIAL.

FT TRANSMEM 148 168 POTENTIAL.

FT TRANSMEM 184 204 POTENTIAL.

FT TRANSMEM 210 230 POTENTIAL.

FT TRANSMEM 234 254 POTENTIAL.

FT TRANSMEM 259 279 POTENTIAL.

SQ SEQUENCE 282 AA; 30768 MW; 9091527BA9749B99 CRC64;

Query Match 79.5%; Score 35; DB 1; Length 282;

Best Local Similarity 87.5%; Pred. No. 4.5;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 STAVFFAG 8

Db 245 STGVFFAG 252

RESULT 3

GSPJ_KLEPN ID GSPJ_KLEPN STANDARD; PRT; 198 AA.

DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Probable protease htpX homolog (EC 3.4.24.-).
 GN HTPX OR RS00074 OR RS02252
 OS Ralstonia solanacearum (Pseudomonas solanacearum).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Burkholderiaceae; Ralstonia.
 OX NCBI_TaxID=305;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GM11000;
 RX MEDLINE=21681879; PubMed=11923852;
 RA Salanoubat M., Genin S., Ariguenave F., Gouzy J., Mangerot S.,
 RA Arlat M., Billault A., Brotier P., Camus J.C., Catolico L.,
 RA Chandler M., Choigne N., Claudel-Renard C., Cunnac S., Demange N.,
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
 RA Sigulier P., Thebault P., Whalen M., Winkler P., Levy M.,
 RA Weissbach J., Boucher C.A.;
 RT "Genome sequence of the plant pathogen Ralstonia solanacearum";
 RL Nature 415:497-502(2002).
 CC -|- COPACOR: Binds 1 zinc ion per subunit (By similarity).
 CC -|- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M48.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AL646057; CAD13602.1; -
 DR HAMAP; MF_00188; -; 1.
 DR InterPro; IPR001915; Peptidase M48.
 DR InterPro; IPR006025; Zn_Mpeptidase.
 DR Pfam; PF01435; Peptidase M48; 1.
 DR PROSITE; PS00142; ZINC PROTEASE; 1.
 DR Transmembrane; Hydrolase; Metalloprotease; Zinc; Complete proteome.
 KW TRANSMEM 7
 FT TRANSNEM 30 POTENTIAL.
 FT TRANSNEM 49 POTENTIAL.
 FT TRANSNEM 142 POTENTIAL.
 FT TRANSNEM 174 POTENTIAL.
 FT TRANSNEM 193 POTENTIAL.
 FT ACT_SITE 132 133
 FT METAL 131 131
 FT METAL 135 135
 FT METAL 135 135
 SQ SEQUENCE 286 AA; 30702 MW; D7E2P2C1593380D0 CRC64;
 Query Match 70.5%; Score 31; DB 1; Length 286;
 Best Local Similarity 85.7%; Pred. No. 31;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 3 AVFPAGR 9
 DB 161 AVFPGR 167
 RESULT 7
 ID PUR2_RHIME STANDARD; PRT; 423 AA.
 AC Q92RL0;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Phosphoribosylamine--glycine ligase (EC 6.3.4.13) (GARS) (Glycinamide
 DE ribonucleotide synthetase)
 GN PURD OR R00858 OR SMC00993.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=1021;
 RX MEDLINE=21396507; PubMed=11481430;
 RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
 RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
 RA Gohl T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
 RA Pohl T., Portetelle D., Puhler A., Purnelle B., Ramsperger U.,
 RA Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;
 RT "Analysis of the chromosome sequence of the legume symbiont
 RT Sinorhizobium meliloti strain 1021";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
 CC -|- CATALYTIC ACTIVITY: ATP + 5-phospho-D-ribose + glycine = ADP
 CC + phosphate + N(1)-(5-phospho-D-ribose)glycinamide.
 CC -|- PATHWAY: De novo purine biosynthesis; second step.
 CC -|- SIMILARITY: Belongs to the GARS family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AL591785; CAC45430.1; -
 DR HAMAP; MF_00138; -; 1.
 DR InterPro; IPR000115; Gars.
 DR Pfam; PF01071; GARS; 1.
 DR Pfam; PF02842; GARS B; 1.
 DR Pfam; PF02843; GARS C; 1.
 DR Pfam; PF02844; GARS N; 1.
 DR TIGRFAMs; TIGR00877; purD; 1.
 DR PROSITE; PS00194; GARS; 1.
 DR Purine biosynthesis; Ligase; Complete proteome.
 KW SEQUENCE 423 AA; 44324 MW; 5E65E19B606D204B CRC64;
 SQ
 Query Match 70.5%; Score 31; DB 1; Length 423;
 Best Local Similarity 85.7%; Pred. No. 44;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 TAVFPAG 8
 DB 264 TGVFPAG 270
 RESULT 8
 ID ANT3_HUMAN STANDARD; PRT; 464 AA.
 AC P01008; P78439; P78447;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Antithrombin-III precursor (ATIII) (PRO0309).
 GN SERPINC1 OR AT3
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=83143280; PubMed=6298709;
 RA Bock S.C., Wion K.L., Vehar G.A., Lawn R.M.;
 RT "Cloning and expression of the cDNA for human antithrombin III";
 RL Nucleic Acids Res. 10:8113-8125(1982).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=83169777; PubMed=6572945;
 RA Chandra T., Srachhouse R., Kidd V.J., Woo S.L.C.;
 RT "Isolation and sequence characterization of a cDNA clone of human
 RT antithrombin III";
 RL Proc. Natl. Acad. Sci. U.S.A. 80:1845-1848(1983).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93237227; PubMed=8476848;

RA Ofosu F.A., Blajchman M.A.;
 RT "Antithrombin-III-Hamilton: a gene with a point mutation (Guanine to
 RT adenine) in codon 382 causing impaired serine protease reactivity.";
 RL Blood 72:1518-1523(1998).
 RN [24]
 RN VARIANTS GLASGOW AND NORTHWICK-PARK.
 RP MEDLINE=8818689; PubMed=3162733;
 RA Erdjument H., Lamed D.A., Panico M., di Marzo V., Morris H.R.;
 RT "Single amino acid substitutions in the reactive site of antithrombin
 RT leading to thrombosis. Congenital substitution of arginine 393 to
 RT cysteine in antithrombin Northwick Park and to histidine in
 RT antithrombin Glasgow.";
 RL J. Biol. Chem. 263:5589-5593(1988).
 RN [25]
 RN VARIANT CHICAGO.
 RP MEDLINE=89388698; PubMed=2781509;
 RA Erdjument H., Lane D.A., Panico M., di Marzo V., Morris H.R.,
 RA Bauer K., Rosenberg R.D.;
 RT "Antithrombin Chicago, amino acid substitution of arginine 393 to
 RT histidine.";
 RL Thromb. Res. 54:613-619(1989).
 RN [26]
 RN VARIANT ROUEN-4.
 RP MEDLINE=90306344; PubMed=2365065;
 RA Borg J.Y., Brennan S.O., Carrell R.W., George P., Perry D.J., Shaw J.;
 RA "Antithrombin Rouen-IV 24 Arg-->Cys. The amino-terminal contribution
 RT to heparin binding.";

Query Match 70.5%; Score 31; DB 1; Length 464;
 Best Local Similarity 77.8%; Pred. No. 49;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 STAVFFNAG 9
 |||||
 DB 417 STAVVIAGR 425

RESULT 9
 VL2 HPV32 STANDARD; PRT; 476 AA.
 AC P36757;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Minor capsid protein L2.
 GN L2
 OS Human papillomavirus type 32.
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 OC Papillomavirus.
 OX NCBI_TaxID=10612;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94265501; PubMed=8205838;
 RA Delius H., Hofmann B.;
 RA "Primer-directed sequencing of human papillomavirus types.";
 RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X74475; CAA52553.1; -
 DR PIR; S36513; S36513.
 DR InterPro; IPR000784; Late L2.
 DR Pfam; PF00513; late_protein_L2; 1.
 DR Coat protein; late_protein.
 KW SEQUENCE 476 AA; 50815 MW; 713344A2164B01EA CRC64;
 SQ

Query Match 70.5%; Score 31; DB 1; Length 476;

Best Local Similarity 75.0%; Pred. No. 50;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 STAVFFAG 8
 |||||
 DB 49 STGVFFGG 56

RESULT 10
 VL2 HPV49 STANDARD; PRT; 521 AA.
 AC P36762;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Minor capsid protein L2.
 GN L2
 OS Human papillomavirus type 49.
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 OC Papillomavirus.
 OX NCBI_TaxID=10616;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94265501; PubMed=8205838;
 RA Delius H., Hofmann B.;
 RA "Primer-directed sequencing of human papillomavirus types.";
 RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X74480; CAA52583.1; -
 DR PIR; S36571; S36571.
 DR InterPro; IPR000784; Late L2.
 DR Pfam; PF00513; late_protein_L2; 1.
 DR Coat protein; late_protein.
 KW SEQUENCE 521 AA; 57054 MW; 3DFB43327FB2E5C3 CRC64;
 SQ

Query Match 70.5%; Score 31; DB 1; Length 521;
 Best Local Similarity 75.0%; Pred. No. 54;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 STAVFFAG 8
 |||||
 DB 47 STGVFFGG 54

RESULT 11
 SUT2 STYHA STANDARD; PRT; 662 AA.
 ID SUT2 STYHA
 AC P53392;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE High affinity sulphate transporter 2.
 GN ST2.
 OS Stylosanthes hamata (Caribbean stylo).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosoids 1; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
 OC Stylosanthes.
 OX NCBI_TaxID=37660;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Verano; TISSUE=Root;
 RX MEDLINE=96016171; PubMed=7568135;
 RA Smith F.W., Ealing P.M., Hawkesford M.J., Clarkson D.T.;
 RA "Plant members of a family of sulfate transporters reveal functional

```

CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AL445065; CAC11908.1; --
CC HAMAP; MF_00045; -- 1.
CC InterPro; IPR002300; tRNA-synt_1a.
CC InterPro; IPR001412; tRNA-synt_1.
CC Pfam; PF00133; tRNA-synt_1; 1.
CC PROSITE; PS00178; AA TRNA_LIGASE_1; 1.
CC Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
CC Complete proteome. 60 "HIGH" REGION.
CC SITE 50 60
CC SITE 611 615 "KMSKS" REGION.
CC BINDING 614 614 ATP (BY SIMILARITY).
CC SEQUENCE 910 AA; 105027 MW; 17FA01BB946DFOB8 CRC64;
CC -----
CC Query Match 70.5%; Score 31; DB 1; Length 910;
CC Best Local Similarity 62.5%; Pred. No. 92;
CC Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
CC -----
CC QY 2 TAVFFAGR 9
CC DB 222 TAIFFRGK 229
CC -----
CC RESULT 14
CC SYI_HAEN STANDARD; PRT; 941 AA.
CC AC P43824;
CC DT 01-NOV-1995 (Rel. 32, Created)
CC DT 01-NOV-1995 (Rel. 32, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE Tsoleucyl-tRNA synthetase (EC 6.1.1.5) (Isoleucine--tRNA ligase)
CC DE (IleRS)
CC GN ILVES OR HI0962.
CC OS Haemophilus influenzae.
CC OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
CC OC Pasteurellaceae; Haemophilus.
CC OX NCBI_TaxID=727;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=RD / KW20 / ATCC 51907;
CC RX MEDLINE=95350630; PubMed=7542800;
CC RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
CC RA Karlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
CC RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
CC RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
CC RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
CC RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
CC RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
CC RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
CC RA Venter J.C.;
CC RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
CC RT Rd.";
CC RL Science 269:496-512(1995).
CC RN [2]
CC RP CONCEPTUAL TRANSLATION.
CC RA Balroch A.;
CC RL Unpublished observations (AUG-1995).
CC CC -!- CATALYTIC ACTIVITY: ATP + L-isoleucine + tRNA(Ile) = AMP +
CC CC di-phosphate + L-isoleucyl-tRNA(Ile).
CC CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC CC -!- SUBUNIT: Monomer (By similarity).
CC CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC CC -!- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.

```

```

CC -!- CAUTION: THIS IS A CONCEPTUAL TRANSLATION; A STOP CODON HAD TO
CC BE SKIPPED IN POSITION 30 TO PRODUCE THIS ORF.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U32777; -- NOT ANNOTATED_CDS.
CC PIR; S78633; S78633.
CC HSP; P41972; 1QJ3.
CC TIGR; H10962; --
CC InterPro; IPR002300; tRNA-synt_1a.
CC InterPro; IPR001412; tRNA-synt_1.
CC InterPro; IPR002301; tRNA-synt_1le.
CC Pfam; PF00133; tRNA-synt_1; 1.
CC PRINTS; PR00984; TRNASYNTHILE.
CC TIGRams; TIGR00392; ileS; 1.
CC PROSITE; PS00178; AA TRNA_LIGASE_1; 1.
CC Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
CC Metal-binding; Zinc; Complete proteome.
CC SITE 59 69 "HIGH" REGION.
CC SITE 603 607 "KMSKS" REGION.
CC BINDING 606 606 ATP (BY SIMILARITY).
CC SEQUENCE 941 AA; 106525 MW; 7A2063BB3F4219AC CRC64;
CC -----
CC Query Match 70.5%; Score 31; DB 1; Length 941;
CC Best Local Similarity 66.7%; Pred. No. 95;
CC Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
CC -----
CC QY 1 STAVFFAGR 9
CC DB 367 STTEFFAGK 375
CC -----
CC RESULT 15
CC PMPD_CHLMU STANDARD; PRT; 1520 AA.
CC AC O9PLB0;
CC DT 16-OCT-2001 (Rel. 40, Created)
CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
CC DE Probable outer membrane protein pmpp precursor (polymorphic membrane
CC DE protein D).
CC GN PMPD OR TC0197.
CC OS Chlamydia muridarum.
CC OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
CC OX NCBI_TaxID=83560;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=MoPn / Nigg;
CC RX MEDLINE=20150255; PubMed=10684935;
CC RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.P.,
CC RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
CC RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
CC RA Gwin M., Nelson W., deBoy R., Kolonay J., McClarty G., Salzberg S.L.,
CC RA Eisen J., Fraser C.M.;
CC RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
CC RT pneumoniae AR39.";
CC RL Nucleic Acids Res. 28:1397-1406(2000).
CC CC -!- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES)
CC CC (POTENTIAL).
CC CC -!- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

RESULT 2
OS8322 PRELIMINARY; PRT; 220 AA.
AC Q8T322;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE AT24862p.
GN CG9455.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Stapleton M., Brokstein P., Hong L., Agbavani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez S., Guatin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celnikier S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR ENBL; AY089426; AL90164.1;
DR HSPF; P01009; IQLP.
DR FlyBase; FBgn0033113; CG9455.
DR InterPro; IPR000215; Serpin.
DR Pfam; PF00079; serpin; 1.
DR SMART; SM00093; SERPIN; 1.
DR PROSITE; PS00284; SERPIN; 1.
KW Protease inhibitor; Serine protease inhibitor; Serpin.
SQ SEQUENCE 220 AA; 24886 MW; 94C33986D39D6599 CRC64;

Query Match 79.5%; Score 35; DB 5; Length 220;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AVFAGR 9
Db 187 AVFAGR 193

RESULT 3
Q9V989 PRELIMINARY; PRT; 403 AA.
AC Q9V989;
DT 01-MAR-2000 (TREMBLrel. 13, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE CG9455 protein.
GN CG9455.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Berkeley;
RX MEDLINE=20196C06; PubMed=10731132;
RA Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbavani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Balles R.M., Baay A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.V., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

```

```

RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleisschmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Gloddek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.C., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Markovlov G., Malshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mout S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Palaznet K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Celnikier S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busan D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattai B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Swirskas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of Drosophila melanogaster genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celnikier S.E.,
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT "Annotation of Drosophila melanogaster genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celnikier S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR ENBL; AE003790; AAF57407.2;
DR HSP; P05619; HLE
DR FlyBase; FBgn0033113; CG9455.
DR InterPro; IPR000215; Serpin.
DR Pfam; PF00079; serpin; 1.
DR SMART; SM00093; SERPIN; 1.
DR PROSITE; PS00284; SERPIN; 1.
KW Protease inhibitor; Serine protease inhibitor; Serpin.
SQ SEQUENCE 403 AA; 45362 MW; 39A3E9B19B4EAF93 CRC64;

```

GN LW057.
 OS Lumpy skin disease virus (LSDV).
 OC Viruses; deDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC Capripoxvirus.
 OX NCBI_TaxID=59509;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Neethling vaccine LW 1959;
 RA Kara P.D., Afonso C.L., Wallace D.B., Kutish G.F., Stipinovich C.,
 RA Lu Z., Vreede F.T., Rajjaard L.C.F., Zsak A., Viljoen G.J., Rock D.L.;
 RT "Molecular characterization of the South African vaccine strain and
 RT the field isolate of lumpy skin disease virus."
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF409138; AAN02782.1; -
 SQ SEQUENCE 373 AA; 42043 MW; 01BD962A161BC889 CRC64;
 Query Match 75.0%; Score 33; DB 12; Length 373;
 Best Local Similarity 75.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 STAFFPAG 8
 Db 190 STPIFFAG 197
 RESULT 8
 ID Q9FCQ6 PRELIMINARY; PRT; 433 AA.
 AC Q9FCQ6;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE Flagellin (Fragment).
 GN FLIC.
 OS Pseudomonas fluorescens.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=294;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SEW25;
 RA Bellingham N.F., Morgan J.A.W., Saunders J.R., Winstanley C.;
 RT "Flagellin gene sequence phylogeny in the genus Pseudomonas";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ297531; CAC03722.1; -
 DR InterPro; IPR01029; Flagellin_C.
 DR Pfam; PF00700; Flagellin_C; 1.
 DR ProDom; PD000316; Flagellin_C; 1.
 FT NON_TER 1
 FT TER 433
 FT NON_TER 1
 FT TER 433
 SQ SEQUENCE 433 AA; 43888 MW; FEB39AD10690122F CRC64;

Query Match 75.0%; Score 33; DB 2; Length 433;
 Best Local Similarity 66.7%; Pred. No. 1.5e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 STAFFPAG 9
 Db 422 STAILFAG 430

RESULT 9
 ID Q9AWQ8 PRELIMINARY; PRT; 121 AA.
 AC Q9AWQ8;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE P0426B06.12 protein.
 GN P0426B06.12.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Ehrhartoideae; Oryzeae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
 RT clone: P0426D06.";
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF002953; BAB21232.1; -
 DR Gramine; Q9AWQ8;
 SQ SEQUENCE 121 AA; 13022 MW; D7A38817C0591148 CRC64;
 Query Match 72.7%; Score 32; DB 10; Length 121;
 Best Local Similarity 75.0%; Pred. No. 71;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 STAFFPAG 8
 Db 36 SSAFFSG 43
 RESULT 10
 ID Q8EZIO PRELIMINARY; PRT; 199 AA.
 AC Q8EZIO;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein.
 GN LA3873.
 OS Leptospira interrogans.
 OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
 OX NCBI_TaxID=173;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
 RA Ren S.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE011542; AAN51071.1; -
 DR Hypothetical protein; Complete proteome.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 199 AA; 22136 MW; 30A0F2F3D76C974 CRC64;
 Query Match 72.7%; Score 32; DB 16; Length 199;
 Best Local Similarity 85.7%; Pred. No. 1.1e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 TAVFFAG 8
 Db 174 TAVFFSG 180

RESULT 11
 ID Q9EXA8 PRELIMINARY; PRT; 221 AA.
 AC Q9EXA8;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Cytochrome C-type biogenesis protein, heme exporter protein B.
 GN DR0407.
 OS Deinococcus radiodurans.
 OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
 OC Deinococcaceae; Deinococcus.
 OX NCBI_TaxID=1299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=20036896; PubMed=10567266;
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
 RA Dodson R.J., Hain D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
 RA Morfat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
 RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,

DE Putative LeOPT1-oligopeptide transporter.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Wing R.A., Frisch D., Presting G., Wood T., Yu Y., Soderlund C.,
RA Kim H., Rambo T., Henry D., Simmons J.;
RT "Rice Genomic Sequence."
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC079685; AAK52575.1; -.
DR Gramene; Q94LJ1; -.
DR InterPro; IPR000109; PTR2.
DR Pfam; PF00854; PTR2; 1.
SQ SEQUENCE 566 AA; 62274 MW; F540ED68EF531960 CRC64;

Query Match 72.7%; Score 32; DB 10; Length 566;
Best Local Similarity 85.7%; Pred. No. 3.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 AVFFAGR 9
Db 238 ALFFAGR 244

Search completed: December 11, 2003, 18:28:47
Job time : 34.333 secs

DR EMBL; AC008261; AAF26166.1; -.
DR InterPro: IPR003657; WRKY_2
DR Pfam: PF03106; WRKY_2
DR PROSITE; PS00611; WRKY_2
KW Transcription regulation; DNA-binding; Nuclear protein; Repeat.
FT DNA_BIND 161 225 WKY 1.
FT DNA_BIND 300 365 WKY 2.
FT DOMAIN 150 154 POLY-ASN.
FT DOMAIN 257 270 POLY-ASP.
FT CONFLICT 81
FT SLL (IN REF. 2).
FT SEQUENCE 423 AA; 47121 MW; EF6C2F1BF3F16B15 CRC64;
SQ

Query Match 75.0%; Score 33; DB 1; Length 423;
Best Local Similarity 77.8%; Pred.No. 26;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 STEVEGAGR 9
DB 236 STEVGAGQ 244

RESULT 3
E2B2 PYRHO
ID _E2B2 PYRHO STANDARD; PRT; 324 AA.
AC O57947;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative translation initiation factor eIF-2B subunit 2 (eIF-2B GDP-GTP exchange factor).
GN PHO208.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=53953;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=98344137; PubMed=9679194;
RX Kawaiabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y., Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y., Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y., Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K., Masuchi Y., Shizuwa H., Kikuchi H.;
RA "Complete sequence and gene organization of the genome of a hyperthermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
RL Nucleic Acids Res. 5:55-76(1998).
CC -!- FUNCTION: CATALYZES THE EXCHANGE OF INITIATION FACTOR 2-BOUND GDP FOR GTP (BY SIMILARITY).
CC -!- SUBUNIT: COMPLEX OF TWO DIFFERENT SUBUNITS (POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE EIF-2B ALPHA/BETA/Delta SUBUNITS FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).

EMBL; AP000001; BAA29277.1; -.
DR PIR; F71243; F71243.
DR InterPro: IPR005250; AIF-2BII_fam.
DR InterPro: IPR006649; IF-2B.
DR Pfam: PF01008; IF-2B; 1.
DR TIGRFAMs; TIGR00511; aif-2BII fam; 1.
DR TIGRFAMs; TIGR00524; eif-2B-rel; 1.
KW Hypothetical protein; Initiation factor; Protein biosynthesis;
KW Complete proteome.
SQ SEQUENCE 324 AA; 36406 MW; A78FC28ACD5D2AB CRC64;

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; AE009633; AAL53185.1; -
 CC FIR; AF3502; AF3502.
 CC HAMAP; MF_00283; -; 1.
 CC InterPro; IPR005146; B3_4.
 CC InterPro; IPR005147; B5.
 CC InterPro; IPR005121; Fdx-AntiCB.
 CC InterPro; IPR004532; Phet_bact.
 CC InterPro; IPR002547; tRNA_bind.
 CC Pfam; PF03483; B3_4; 1.
 CC Pfam; PF03484; B5; 1.
 CC Pfam; PF03147; FDX-ACB; 1.
 CC Pfam; PF01588; tRNA_bind; 1.
 CC TIGRFAMs; TIGR00472; phet_bact; 1.
 CC PROSITE; PS50886; TRBD; 1.
 CC Aminocacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 CC Metal-binding; Magnesium; RNA-binding; tRNA-binding;
 CC Complete proteome.
 CC DOMAIN 38 148 TRNA-BINDING.
 CC METAL 454 454 MAGNESIUM (BY SIMILARITY).
 CC METAL 460 460 MAGNESIUM (VIA CARBONYL OXYGEN) (BY
 CC SIMILARITY).
 CC METAL 463 463 MAGNESIUM (BY SIMILARITY).
 CC METAL 464 464 MAGNESIUM (BY SIMILARITY).
 CC SEQUENCE 804 AA; 85988 MW; CD1063F9B05E70C1 CRC64;

Query Match 72.7%; Score 32; DB 1; Length 804;
 Best Local Similarity 66.7%; Pred. No. 82;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 STEVEGAGR 9
 : : : : :
 Db 597 TAKVEGAGR 605

RESULT 7
 ID SYFB BRUSU STANDARD; PRT; 804 AA.
 AC Q8FXA4;
 DT 15-SEP-2003 (Rel. 42, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Phenylalanyl-tRNA synthetase beta chain (EC 6.1.1.20)
 DE (Phenylalanine--tRNA ligase beta chain) (PHERS).
 GN PHET OR BR2123.
 OS Brucella suis.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 CC Brucellaceae; Brucella.
 OX NCBI_TaxID=29461;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1330 / Biovar 1;
 RX MEDLINE=2247741; PubMed=12271122;
 RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,
 RA Read T.D., Dodson R.J., Umayam L., Brinkac L.M., Beanan M.J.,
 RA Daugherty S.C., Deboy R.T., Durkin A.S., Kolonay J.F., Madupu R.,
 RA Nelson W.C., Ayodeji B., Knaul M., Shetty J., Malek J., Van Aken S.E.,
 RA Riedmuller S., Tettelin H., Gill S.R., White O., Salzberg S.L.,
 RA Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M., Fraser C.M.;
 RT "The Brucella suis genome reveals fundamental similarities between
 RT animal and plant pathogens and symbionts."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153 (2002).
 CC -/- CATALYTIC ACTIVITY: ATP + L-phenylalanine + tRNA (Phe) = AMP +
 CC diphosphate + L-phenylalanyl-tRNA (Phe).
 CC -/- COFACTOR: Binds 2 magnesium ions per tetramer (By similarity).
 CC -/- SUBUNIT: Tetramer of two alpha and two beta chains (By
 CC similarity).

CC -/- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -/- SIMILARITY: Belongs to the phenylalanyl-tRNA synthetase beta chain
 CC family. Subfamily 1.
 CC -/- SIMILARITY: Contains 1 tRNA-binding domain.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; AE014499; AAN31013.1; -
 CC TIGR; BR2123; -
 CC HAMAP; MF_00283; -; 1.
 CC InterPro; IPR005146; B3_4.
 CC InterPro; IPR005147; B5.
 CC InterPro; IPR005121; Fdx-AntiCB.
 CC InterPro; IPR004532; Phet_bact.
 CC InterPro; IPR002547; tRNA_bind.
 CC Pfam; PF03483; B3_4; 1.
 CC Pfam; PF03484; B5; 1.
 CC Pfam; PF03147; FDX-ACB; 1.
 CC Pfam; PF01588; tRNA_bind; 1.
 CC TIGRFAMs; TIGR00472; phet_bact; 1.
 CC PROSITE; PS50886; TRBD; 1.
 CC Aminocacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 CC Metal-binding; Magnesium; RNA-binding; tRNA-binding;
 CC Complete proteome.
 CC DOMAIN 38 148 TRNA-BINDING.
 CC METAL 454 454 MAGNESIUM (BY SIMILARITY).
 CC METAL 460 460 MAGNESIUM (VIA CARBONYL OXYGEN) (BY
 CC SIMILARITY).
 CC METAL 463 463 MAGNESIUM (BY SIMILARITY).
 CC METAL 464 464 MAGNESIUM (BY SIMILARITY).
 CC SEQUENCE 804 AA; 66857F8268E875D CRC64;

Query Match 72.7%; Score 32; DB 1; Length 804;
 Best Local Similarity 66.7%; Pred. No. 82;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 STEVEGAGR 9
 : : : : :
 Db 597 TAKVEGAGR 605

RESULT 8
 ID CLJA_BACTU STANDARD; PRT; 1167 AA.
 AC Q45738;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Pesticidal crystal protein cryIIJa (Insecticidal delta-endotoxin
 DE CryIIJa) (Crytalline entomocidal protoxin) (133 kDa crystal protein).
 GN CRYIIJA OR CRYIIJA OR CRYET4.
 OS Bacillus thuringiensis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1428;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL B-21110 / EG8947;
 RA Donovan W.P., Tan Y., Jany C.S., Gonzalez J.M. Jr.;
 RA "Bacillus thuringiensis cryet4 and cryet5 toxin genes and proteins
 RA toxic to lepidopteran insects.";
 RT Patent number US5322687, 21-JUN-1994.
 RL
 CC -/- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
 CC EPITHELIAL CELLS OF MANY LEPIDOPTERAN LARVAE.
 CC -/- DEVELOPMENTAL STAGE: THE CRYSTAL PROTEIN IS PRODUCED DURING
 CC SPOROGATION AND IS ACCUMULATED BOTH AS AN INCLUSION AND AS PART
 CC OF THE SPORE COAT.
 CC -/- MISCELLANEOUS: TOXIC SEGMENT OF THE PROTEIN IS LOCATED IN THE

Fri Dec 12 13:22:02 2003

us-10-014-658-10.rsp

```

RESULT 13
ZPR1_HUMAN
ID ZPR1_HUMAN STANDARD; PRT; 459 AA.
AC 075312;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Zinc-finger protein ZPR1 (zinc finger protein 259).
GN ZNF259 OR ZPR1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RA Galcheva-Gargova Z., Konstantinov K.N., Gangwani L., Mikrut M.,
RA Purohit P., Theroux S.J., Enoch T., Davis R.J.;
RT "Translational regulation by the ZPR1 signal transduction pathway.";
RT Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.
RN [2]
SEQUENCE FROM N.A.
RA Galcheva-Gargova Z., Konstantinov K.N., Gangwani L., Mikrut M.,
RA Purohit P., Theroux S.J., Enoch T., Davis R.J.;
RT "Translational regulation by the ZPR1 signal transduction pathway.";
RT Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.
RN [3]
SUBCELLULAR LOCATION.
MEDLINE=98437195; PubMed=9763455;
RA Galcheva-Gargova Z., Gangwani L., Konstantinov K.N., Mikrut M.,
RA Theroux S.J., Enoch T., Davis R.J.;
RT "The cytoplasmic zinc finger protein ZPR1 accumulates in the nucleolus
of proliferating cells.";
RL Mol. Biol. Cell 9:2963-2971(1998).
CC -!- FUNCTION: MAY BE A SIGNALING MOLECULE THAT COMMUNICATES MITOGENIC
CC SIGNALS FROM THE CYTOPLASM TO THE NUCLEUS.
CC -!- SUBUNIT: BINDS TO THE EGF AND PDGF RECEPTORS. BINDS TO THE
CC ELONGATION FACTOR 1-ALPHA (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC; TRANSLOCATE TO THE NUCLEUS
CC AFTER TREATMENT WITH MITOGENS.
CC -!- SIMILARITY: BELONGS TO THE ZPR1 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF019767; AAC33514.1; -
CC EMBL; BC004256; AAH04256.1; -
CC EMBL; EC012182; AAH12182.1; -
CC EMBL; BC017349; AAH17349.1; -
CC EMBL; BC017380; AAH17380.1; -
CC Genew; HGNC:13051; ZNF259.

Query Match 70.5%; Score 31; DB 1; Length 459;
Best Local Similarity 55.8%; Pred. No. 75;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 STEVEGAGR 9
DB 87 NTEIQSAGR 95

RESULT 14
ZPR1_MOUSE
ID ZPR1_MOUSE STANDARD; PRT; 459 AA.
AC Q62384;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Zinc-finger protein ZPR1 (zinc finger protein 259).
GN ZNF259 OR ZPR1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RA Galcheva-Gargova Z., Konstantinov K.N., Wu I.-H., Klier F.G.,
RA Barrett T., Davis R.J.;
RT "Binding of zinc finger protein ZPR1 to the epidermal growth
factor receptor.";
RL Science 272:1797-1802(1996).
RN [2]
SUBCELLULAR LOCATION.
MEDLINE=98437195; PubMed=9763455;
RA Galcheva-Gargova Z., Gangwani L., Konstantinov K.N., Mikrut M.,
RA Theroux S.J., Enoch T., Davis R.J.;
RT "The cytoplasmic zinc finger protein ZPR1 accumulates in the nucleolus
of proliferating cells.";
RL Mol. Biol. Cell 9:2963-2971(1998).
CC -!- FUNCTION: MAY BE A SIGNALING MOLECULE THAT COMMUNICATES MITOGENIC
CC SIGNALS FROM THE CYTOPLASM TO THE NUCLEUS.
CC -!- SUBUNIT: BINDS TO THE EGF AND PDGF RECEPTORS. BINDS TO THE
CC ELONGATION FACTOR 1-ALPHA.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC; TRANSLOCATE TO THE NUCLEUS
CC AFTER TREATMENT WITH MITOGENS.
CC -!- SIMILARITY: BELONGS TO THE ZPR1 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U41287; AAC52662.1; -
CC MGD; MGI:1330262; Zfp259.
CC GO; GO:0005737; C:cytoplasm; IDA.
CC GO; GO:0005634; C:nucleus; IDA.
CC GO; GO:0005515; F:protein binding activity; IPI.

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2003, 18:03:30 ; Search time 20.3333 Seconds
(without alignments)
114.220 Million cell updates/sec

Title: US-10-014-658-10

Perfect score: 44

Sequence: 1 STEVEGAGR 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: SP_ARCHAEA.*
- 2: SP_BACTERIA.*
- 3: SP_FUNGI.*
- 4: SP_HUMAN.*
- 5: SP_INVERTEBRATE.*
- 6: SP_MAMMAL.*
- 7: SP_MHC.*
- 8: SP_ORGANALLE.*
- 9: SP_PHAGE.*
- 10: SP_PLANT.*
- 11: SP_RODENT.*
- 12: SP_VIRUS.*
- 13: SP_VERTEBRATE.*
- 14: SP_UNCLASSIFIED.*
- 15: SP_VIRUS.*
- 16: SP_BACTERIAP.*
- 17: SP_ARCHAEP.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	81.8	253	Q8EMH7	Q8EMH7 oceanobacil
2	36	81.8	674	O50431	O50431 mycobacteri
3	36	81.8	711	Q8TH00	Q8TH00 methanopyru
4	36	81.8	824	Q8P357	Q8P357 xanthomonas
5	34	77.3	135	Q8NS07	Q8NS07 corynebacte
6	33	75.0	157	Q8XMU4	Q8XMU4 clostridium
7	33	75.0	259	Q8KZ62	Q8KZ62 streptomyce
8	33	75.0	263	Q8XAL6	Q8XAL6 streptomyce
9	33	75.0	294	Q9VPL4	Q9VPL4 drosophila
10	33	75.0	316	Q8PP70	Q8PP70 xanthomonas
11	33	75.0	318	Q8PCHO	Q8PCHO xanthomonas
12	33	75.0	339	Q41113	Q41113 phaseolus v
13	33	75.0	345	Q8FNBS	Q8FNBS corynebacte
14	33	75.0	2174	Q9GQRO	Q9GQRO drosophila
15	33	75.0	3060	Q9VAV4	Q9VAV4 drosophila
16	32	72.7	120	Q8TAS2	Q8TAS2 homo sapien

17	32	72.7	274	10	Q8RVE3	Q8RVE3 oryza sativ
18	32	72.7	291	16	Q8ZSD6	Q8ZSD6 anabaena sp
19	32	72.7	305	16	Q8ZN23	Q8ZN23 salmorella
20	32	72.7	305	16	Q8Z410	Q8Z410 salmorella
21	32	72.7	317	2	Q8Z117	Q8Z117 streptomyce
22	32	72.7	318	16	Q8L026	Q8L026 streptomyce
23	32	72.7	360	10	Q9FZJ3	Q9FZJ3 arabidopsis
24	32	72.7	423	16	Q8FMX4	Q8FMX4 corynebacte
25	32	72.7	445	16	Q9FZJ0	Q9FZJ0 streptomyce
26	32	72.7	456	16	Q8E5N5	Q8E5N5 streptococc
27	32	72.7	456	16	Q8DZV4	Q8DZV4 streptococc
28	32	72.7	494	16	Q8XON5	Q8XON5 ralstonia s
29	32	72.7	541	16	Q8RFJ3	Q8RFJ3 fusobacteri
30	32	72.7	553	10	Q04518	Q04518 arabidopsis
31	32	72.7	610	16	Q8YM91	Q8YM91 anabaena sp
32	32	72.7	802	10	Q8RXV4	Q8RXV4 arabidopsis
33	32	72.7	804	16	Q8YE74	Q8YE74 brucella me
34	32	72.7	804	16	Q8PFX4	Q8PFX4 brucella su
35	32	72.7	837	10	Q80590	Q80590 arabidopsis
36	31	70.5	110	16	Q8YM01	Q8YM01 anabaena sp
37	31	70.5	178	16	Q8Z5J3	Q8Z5J3 mycobacteri
38	31	70.5	189	16	Q8PM02	Q8PM02 xanthomonas
39	31	70.5	193	4	Q8NDRO	Q8NDRO homo sapien
40	31	70.5	253	2	Q8GFC3	Q8GFC3 photorhabdu
41	31	70.5	259	16	Q8HWC9	Q8HWC9 pseudomonas
42	31	70.5	260	10	Q8LCV9	Q8LCV9 arabidopsis
43	31	70.5	279	12	Q8ENS7	Q8ENS7 cucumber mo
44	31	70.5	282	10	Q8H2L6	Q8H2L6 oryza sativ
45	31	70.5	286	6	Q8HX98	Q8HX98 pan troglod

ALIGNMENTS

RESULT 1
Q8EMH7 PRELIMINARY; PRT; 253 AA.

AC Q8EMH7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE 4-hydroxyphenylacetate degradation bifunctional isomerase ;
DE decarboxylase (divided with Q82865 and Q82866) (EC 4.1.1.-).
GN Q82865.
OS Oceanobacillus iheyensis.
OC Bacteria; Firmicutes; Bacillales; Oceanobacillus.
OX NCBI_TaxID=182710;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=HTE831 / DSM 14371 / JCM 11309;
RC MEDLINE=2220767; PubMed=12235376;
RA Takami H., Takaki Y., Uchiyama I.;
RT "Genome sequence of Oceanobacillus iheyensis isolated from the Iheya
RT Ridge and its unexpected adaptive capabilities to extreme
RT environments."
RL Nucleic Acids Res. 30:3927-3935(2002).
DR EMBL; AP004602; BAC14821.1; -;
KW Isomerase; lyase; Complete proteome.
SQ SEQUENCE 253 AA; 28312 MW; 6404088BC1D4AF33 CRC64;

Query Match 81.8%; Score 36; DB 16; Length 253;
Best Local Similarity 87.5%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TEVEGAGR 9
Db 232 TEVEGVGR 239

RESULT 2
O50431 PRELIMINARY; PRT; 674 AA.
ID O50431
AC O50431;

Fri Dec 12 13:22:02 2003

```

RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Seeger S., Harris D.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA James K.D., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapante D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.;
RL Mol. Microbiol. 21:77-96(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Batsman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Segar K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren J., Wierozurek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL939129; CAB45588.1; -
DR InterPro; IPR004360; Gly_bleo_diox.
DR Pfam; PF00903; Glyoxalase; 2.
DR Complete Proteome.
RX SEQUENCE 263 AA; 27276 MW; F9A9F43F8E4976F9 CRC64;
SQ SEQUENCE 263 AA; 27276 MW; F9A9F43F8E4976F9 CRC64;

Query Match 75.0%; Score 33; DB 16; Length 263;
Best Local Similarity 55.6%; Pred. No. 1.6e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 STEVEGAGR 9
Db 233 ATDIEGVGR 241

RESULT 9
Q9VPL4 PRELIMINARY; PRT; 294 AA.
AC Q9VPL4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 20, Last annotation update)
DE CG11617 protein (RE08174p).
GN CG11617.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amaralides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

```

```

RA Abril J.P., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Balaw R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahike C., Davenport L.B., Davies F.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,
RA Fessler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodde A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kaulush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Markulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirski R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers B.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarini H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RX SEQUENCE 294 AA; 34029 MW; F32A0CCFEE1DDAC CRC64;
SQ SEQUENCE 294 AA; 34029 MW; F32A0CCFEE1DDAC CRC64;

Query Match 75.0%; Score 33; DB 5; Length 294;
Best Local Similarity 66.7%; Pred. No. 1.8e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 STEVEGAGR 9
Db 254 TTDVSGAGR 262

RESULT 10
Q9PF70 PRELIMINARY; PRT; 316 AA.
AC Q9PF70;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-WAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Ribokinase
GN RESK OR XAC0818.
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;

```

01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
 DE Putative sodium-dependent transporter.
 GN CE2229.
 OS Corynebacterium efficiens.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
 OX NCB1_TaxID=152794;
 [1]
 RN
 RP
 RC STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;
 RA Kawarayashi Y., Yamazaki J., Hino Y., Kikuchi H., Nakamura Y.,
 RA Ikeo K., Suzuki M., Mashima J., Itoh T., Yamagishi A., Nishio Y.,
 RA Usuda Y., Sugimoto S.;
 RT "The entire genomic sequence of Corynebacterium efficiens YS-314.";
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBSJ databases.
 DR EMBL; AF005221; BAC19039.1; -;
 SK Complete proteome.
 SQ SEQUENCE 345 AA; 36190 MW; 07293DCE94512812 CRC64;
 Query Match 75.0%; Score 33; DB 16; Length 345;
 Best Local Similarity 75.0%; Pred. No. 2.1e+02;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 STEVEGAG 8
 Db 179 ATEVDGAG 186
 RESULT 14
 Q9GQRO PRELIMINARY; PRT; 2174 AA.
 AC Q9GQRO;
 DT 01-MAR-2001 (T-EMBLrel. 16, Created)
 DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
 DE Extracellular matrix protein papilin precursor.
 DE PPN OR CG1540 OR CG18436.
 GN Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCB1_TaxID=7227;
 [1]
 RN
 RP
 RC SEQUENCE FROM N.A.
 RC STRAIN=cp cn bw;
 RX MEDLINE=20530459; PubMed=11076767;
 RA Kramerova I.A., Kawaguchi N., Nelson R.E., Fessler L.I., Chen Y.,
 RA Kramerov A.A., Kusche-Gullberg M., Kramer J.M., Ackley B.D.,
 RA Sieron A.L., Prockop D.J., Fessler J.H.;
 RT "Papilin in development: a pericellular protein with a homology to the
 ADAMTS metalloproteinases";
 RL Development 127:5475-5485(2000).
 DR EMBL; AF205357; AAG37995.1; -;
 DR HSP; P12111; 2KNT.
 DR FlyBase; FBgn003137; Ppn.
 DR InterPro; IPR006209; EGF like.
 DR InterPro; IPR003598; Ig C2.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR002223; Kunitz_BPTI.
 DR InterPro; IPR000884; TSF1.
 DR InterPro; IPR002221; WAP.
 DR Pfam; PF00047; Ig; 2.
 DR Pfam; PF00014; Kunitz_BPTI; 3.
 DR Pfam; PF00090; tsp.1; 5.
 DR Pfam; PF00095; wap_1.
 DR PRINTS; PR00759; BASICPTASE.
 DR ProDom; PDOC00222; Kunitz_BPTI; 3.
 DR SMART; SMO0408; IGC2; 2.
 DR SMART; SMO0131; KU; 3.
 DR SMART; SMO0209; TSF1; 7.
 DR SMART; SMO0217; WAP; 1.
 DR PROSITE; PS00317; 4_DISULFIDE_CORE; 1.

DR PROSITE; PS00280; BPTI_KUNITZ_1; 3.
 DR PROSITE; PS00279; BPTI_KUNITZ_2; 3.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS00092; TSF1; 5.
 KW Immunoglobulin domain; Matrix protein; Protease inhibitor;
 KW Serine protease inhibitor; Signal.
 FT SIGNAL 1 26 POTENTIAL.
 SQ SEQUENCE 2174 AA; 231936 MW; 038F707952623120 CRC64;
 Query Match 75.0%; Score 33; DB 5; Length 2174;
 Best Local Similarity 55.6%; Pred. No. 1.6e+03;
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
 QY 1 STEVEGAG 9
 Db 1349 TTEIEGSGQ 1357
 RESULT 15
 Q9VAV4 PRELIMINARY; PRT; 3060 AA.
 ID Q9VAV4;
 AC Q9VAV4; Q9VAV3;
 DT 01-MAY-2000 (T-EMBLrel. 13, Created)
 DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
 DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
 DE CG1540 protein.
 GN PPN OR CG1540 OR CG18436.
 OS Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCB1_TaxID=7227;
 [1]
 RN
 RP
 RC SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
 Brandon R.C., Rogers Y.-H.C., Blazer V.G., Champagne M., Pfeiffer B.D.,
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 ABail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 BAilew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 Beeson K.Y., Bencs P.V., Bernan B.P., Bhandari D., Bolshakov S.,
 Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
 de Fabros B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 Glodek A., Gong F., Gorell J.H., Gu Z., Guan P., Harris M.,
 Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 Jimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
 Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,
 Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 SVirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
 Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."

```

XX PF 12-MAY-1999; 99WO-US10549.
XX PF
XX PR 12-MAY-1998; 98US-0085197.
XX PR 05-MAY-1999; 99US-0085197.
XX PF
XX PA (BOCK/) BOCK S C.
XX PA (PICA/) PICARD V.
XX PA (ZEN/) ZENDEHROUH P.
XX PF
XX PI Bock SC, Picard V, Zendehtrouh P;
XX PF
XX DR WPI; 2000-116274/10.
XX PF
XX PT New modified human antithrombin III compounds, used for treating e.g.
XX PT sepsis, trauma, acute respiratory distress syndrome, restenosis,
XX PT thrombosis, thromboembolism or stroke -
XX PF
XX PS Claim 13; Page 57; 75pp; English.
XX PF
XX CC The present sequence is from an antithrombin III (ATIII) variant, Aa
XX CC derived from human ATIII-N135A cDNA insert of the pBlueBac baculovirus
XX CC expression construct and comprises residues 385-393. The variant has
XX CC improved resistance to elastase and Igg-activated neutrophils while
XX CC retains anti-thrombin and anti-factor Xa activities. It may be
XX CC expressed as glycoforms with enhanced heparin affinity which target the
XX CC blood vessel wall more efficiently than ATIIIs with normal heparin
XX CC affinity. The modified ATIIIs can be used to treat thrombin activation-
XX CC related pathological symptoms due to sepsis, trauma, acute
XX CC respiratory distress syndrome, restenosis, thrombosis, thromboembolism
XX CC and stroke. It can also be used to reduce the risk of reocclusion
XX CC and restenosis in percutaneous transluminal coronary angioplasty,
XX CC thrombosis associated with surgery, ischaemia/reperfusion injury, and
XX CC coagulation abnormalities in cancer or surgical patients.
XX PF
XX SQ Sequence 9 AA;
XX PF
Query Match 100.0%; Score 47; DB 21; Length 9;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STREGGGGR 9
| | | | |
DB 1 STREGGGGR 9

RESULT 2
AAO08171
ID AAO08171 standard; Protein; 53 AA.
AC AAO08171;
XX PF
XX DT 06-NOV-2001 (first entry)
XX PF
XX DE Human polypeptide SEQ ID NO 22063.
XX PF
XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX KW tissue growth factor; immunomodulatory; cancer; leukaemia;
XX KW nervous system disorders; arthritis; inflammation.
XX PF
XX OS Homo sapiens.
XX PF
XX PN WO200164835-A2.
XX PF
XX PD 07-SEP-2001.
XX PF
XX PF 26-FEB-2001; 2001WO-US04927.
XX PF
XX PR 28-FEB-2000; 2000US-0515126.
XX PR 18-MAY-2000; 2000US-0577409.
XX PF
XX PA (HYSE-) HYSEQ INC.
XX PF

XX PF Tang YT, Liu C, Drmanac RT;
XX PF
XX DR WPI; 2001-514838/56.
XX DR N-PSDB; AAI88102.
XX PF
XX PT Isolated nucleic acids and polypeptides, useful for preventing
XX PT diagnosing and treating e.g. leukaemia, inflammation and immune
XX PT disorders -
XX PF
XX PS Claim 20; SEQ ID NO 22063; 1399pp + Sequence Listing; English.
XX PF
XX CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
XX CC the encoded proteins (AAO0010-AAO13910) that exhibit activity relating to
XX CC cytokine, cell proliferation or cell differentiation or which may induce
XX CC production of other cytokines in other cell populations. The
XX CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX CC peptide therapy. The polypeptides have various cytokine-like activities,
XX CC e.g. stem cell growth factor activity, haematopoiesis regulating
XX CC activity, tissue growth factor activity, immunomodulatory activity and/or
XX CC activin/inhibin activity and may be useful in the diagnosis and/or
XX CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX CC inflammation.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX PF
XX SQ Sequence 53 AA;
XX PF
Query Match 91.5%; Score 43; DB 22; Length 53;
Best Local Similarity 88.9%; Pred. No. 6.7;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 STREGGGGR 9
| | | | |
DB 25 STVEGGGGR 33

RESULT 3
AAV13023
ID AAV13023 standard; Protein; 95 AA.
XX PF
XX AC AAV13023;
XX PF
XX DT 22-JUN-1999 (first entry)
XX PF
XX DE Human secreted protein encoded by 5' EST SEQ ID NO: 37.
XX PF
XX KW Human; secreted protein; EST; expressed sequence tag; diagnosis;
XX KW forensic; gene therapy; chromosome mapping; signal peptide;
XX KW upstream regulatory sequence; cytokine activity; cell proliferation;
XX KW differentiation; haematopoiesis regulation; tissue growth regulation;
XX KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
XX KW thrombolytic; anti-inflammatory; tumour inhibition.
XX PF
XX OS Homo sapiens.
XX PF
XX PN WO9906552-A2.
XX PF
XX PD 11-FEB-1999.
XX PF
XX PF 31-JUL-1998; 98WO-IB01236.
XX PF
XX PR 01-AUG-1997; 97US-0905223.
XX PF
XX PA (GEST ) GENSET.
XX PF
XX PI Duclert A, Dumas Milne Edwards J, Lacroix B;
XX PF
XX DR WPI; 1999-153782/13.
XX DR N-PSDB; AAX51823.
XX PF
XX PT New isolated brain-derived nucleic acids - used to develop products
XX PF

```

polynucleotides may be used for detection of prostate cancer, chromosome identification, as chromosome markers, and for numerous other diagnostic or research purposes. The prostate cancer antigens may be used to treat disorders such as neural, immune, muscular, reproductive, gastrointestinal, pulmonary, cardiovascular, renal, and proliferative disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to AAB57303 represent sequences used in the exemplification of the present invention.

XX Sequence 174 AA;
SQ

Query Match 83.0%; Score 39; DB 21; Length 174;
Best Local Similarity 87.5%; Pred. No. 86;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TAEGGGGR 9
| | | | |
DB 7 TTGGGGGR 14

RESULT 6
ABG91807
ID ABG91807 standard; Protein; 1605 AA.
XX AC ABG91807;
XX DT 29-NOV-2002 (first entry)
XX DE Human intracellular signalling molecule protein INSTIG-9.
XX KW Human; intracellular signalling molecule; INSTIG; atherosclerosis; cell proliferative disease; cancer; autoimmune disease; AIDS; inflammatory disease; acquired immunodeficiency syndrome; allergy; neurological disorder; stroke; Parkinson's disease; epilepsy; gastrointestinal disorder; ulcer; cirrhosis; reproductive disorder; endometriosis; developmental disorder; vesicle trafficking disorder; bacterial infection; viral infection; parasitic infection; protozoal infection.
XX KW
XX OS Homo sapiens.
XX DN W0200263008-A2.
XX PD 15-AUG-2002.
XX PF 07-FEB-2002; 2002WO-US03966.
XX PR 08-FEB-2001; 2001US-267925P.
XX PR 09-MAR-2001; 2001US-274435P.
XX PR 21-MAR-2001; 2001US-277813P.
XX PR 03-APR-2001; 2001US-281328P.
XX PR 15-MAY-2001; 2001US-291195P.
XX PR 16-MAY-2001; 2001US-291550P.
XX PR 25-MAY-2001; 2001US-293591P.
XX PR 01-JUN-2001; 2001US-295349P.
XX PA (INCY-) INCYTE GENOMICS INC.
XX PI Ding L, Warren BA, Elliot VS, Tang YT, Yue H, Burford N, Lee S; Richardson TW, Lal P, Nguyen DB, Yang J, Hafalia AJA, Ison CH; Gururajan R, Baughin MR, Wang YE, Yao MG, Thangavelu K; Swarnakar A, Griffin JA, Forsythe IJ, Emerling BM, Walla NK;
XX DR WPI; 2002-627561/67.
XX DR N-PSDB; ABS67749.
XX PT New human intracellular signalling molecules (INSTIG), useful for diagnosing, treating and preventing diseases or conditions associated with the aberrant INSTIG expression, e.g. cancer, AIDS, atherosclerosis, infections -
XX PS Claim 2; Page 155-158; 195pp; English.

The present invention relates to a new intracellular signalling molecule (INSTIG) polypeptide. The polypeptides and polynucleotides of the invention are useful in diagnosing, treating and preventing diseases or conditions associated with the decreased expression or overexpression of INSTIG, such as cell proliferative diseases (e.g. cancer, atherosclerosis), autoimmune/inflammatory diseases (e.g. AIDS (acquired immunodeficiency syndrome), allergies), neurological disorders (e.g. stroke, Parkinson's disease, epilepsy), gastrointestinal (e.g. ulcer, cirrhosis), reproductive (e.g. endometriosis), developmental, vesicle trafficking disorders, and infections (e.g. bacterial, viral, parasitic, protozoal). These are also useful in assessing the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of INSTIG. The INSTIG or its fragments are useful in screening compounds for effectiveness as agonist or antagonist of the polypeptides, or in altering the expression of the target polynucleotide and compounds that specifically bind to or modulate the activity of the polypeptide. The microarray is useful in monitoring or measuring protein-protein interactions, drug-target interactions, and gene expression profiles. The present amino acid sequence represents a human INSTIG protein of the invention.

XX Sequence 1605 AA;
SQ

Query Match 83.0%; Score 39; DB 23; Length 1605;
Best Local Similarity 87.5%; Pred. No. 7e+02; 0; Gaps 0;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 STAEGGGG 8
| | | | |
DB 4 STADGGGG 11

RESULT 7
AAO12879
ID AAO12879 standard; Protein; 106 AA.
XX AC AAO12879;
XX DT 06-NOV-2001 (first entry)
XX DE Human polypeptide SEQ ID NO 26771.
XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy; vaccine; peptide therapy; stem cell growth factor; haematopoiesis; tissue growth factor; immunomodulatory; cancer; leukaemia; nervous system disorders; arthritis; inflammation.
XX OS Homo sapiens.
XX EN W0200164835-A2.
XX PD 07-SEP-2001.
XX PF 26-FEB-2001; 2001WO-US04927.
XX PR 28-FEB-2000; 2000US-0515126.
XX PR 18-MAY-2000; 2000US-0577409.
XX PR (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Drmanac RT;
XX DR WPI; 2001-514838/56.
XX DR N-PSDB; AAI92810.
XX PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing and treating e.g. leukaemia, inflammation and immune disorders -
XX PS Claim 20; SEQ ID NO 26771; 1399pp + Sequence Listing; English.
XX CC The invention relates to human polynucleotides (AAI93841) and the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to

QY 1 STAEGGGGR 9
Db 67 STSQGGGK 75

RESULT 9
ABG17461
ID ABG17461 standard; Protein; 384 AA.
AC ABG17461;
XX 18-FEB-2002 (first entry)
DT XX
DE XX
XX Novel human diagnostic protein #17452.
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX Homo sapiens.
OS
XX WO200175067-A2.
FN
XX 11-OCT-2001.
PD
XX 30-MAR-2001; 2001WO-US08631.
PF
XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0549167.
XX (HYSE-) HYSEQ INC.
PA
XX Drmanac RT, Liu C, Tang YT;
PI WPI; 2001-639362/73.
XX N-PSDB; AAS81648.
DR
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID No 47820; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (II) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: the sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence. 384 AA;

Query Match 80.9%; Score 38; DB 22; Length 384;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AEGGGGR 9
|||||

PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249254.
PR 17-NOV-2000; 2000US-0249255.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 08-DEC-2000; 2000US-0256719.
PR 08-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI
XX N-PSDB; AAK57510.
DR
XX
XX Nucleic acids encoding human immune/haematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX
XX Claim 11; SEQ ID NO 12322; 3071pp + Sequence Listing; English.
PS
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAK62170 to AAK61921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK67694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK62169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 124 AA;

Query Match 80.9%; Score 38; DB 22; Length 124;
Best Local Similarity 66.7%; Pred. No. 90;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

PR 06-JUL-1999;	99US-0142390.	PR 23-SEP-1999;	99US-0155486.
PR 08-JUL-1999;	99US-0142803.	PR 24-SEP-1999;	99US-0155659.
PR 09-JUL-1999;	99US-0142920.	PR 28-SEP-1999;	99US-0156458.
PR 12-JUL-1999;	99US-0143277.	PR 29-SEP-1999;	99US-0156596.
PR 13-JUL-1999;	99US-0143542.	PR 04-OCT-1999;	99US-0157117.
PR 14-JUL-1999;	99US-0143624.	PR 05-OCT-1999;	99US-0157753.
PR 15-JUL-1999;	99US-0144005.	PR 06-OCT-1999;	99US-0157865.
PR 16-JUL-1999;	99US-0144085.	PR 07-OCT-1999;	99US-0158029.
PR 16-JUL-1999;	99US-0144086.	PR 08-OCT-1999;	99US-0158232.
PR 19-JUL-1999;	99US-0144325.	PR 12-OCT-1999;	99US-0158369.
PR 19-JUL-1999;	99US-0144331.	PR 13-OCT-1999;	99US-0159293.
PR 19-JUL-1999;	99US-0144332.	PR 13-OCT-1999;	99US-0159294.
PR 19-JUL-1999;	99US-0144333.	PR 13-OCT-1999;	99US-0159295.
PR 19-JUL-1999;	99US-0144334.	PR 14-OCT-1999;	99US-0159329.
PR 19-JUL-1999;	99US-0144335.	PR 14-OCT-1999;	99US-0159330.
PR 20-JUL-1999;	99US-0144352.	PR 14-OCT-1999;	99US-0159331.
PR 20-JUL-1999;	99US-0144632.	PR 14-OCT-1999;	99US-0159637.
PR 20-JUL-1999;	99US-0144884.	PR 14-OCT-1999;	99US-0159638.
PR 21-JUL-1999;	99US-0144814.	PR 18-OCT-1999;	99US-0159584.
PR 21-JUL-1999;	99US-0145086.	PR 21-OCT-1999;	99US-0160741.
PR 22-JUL-1999;	99US-0145088.	PR 21-OCT-1999;	99US-0160767.
PR 22-JUL-1999;	99US-0145085.	PR 21-OCT-1999;	99US-0160768.
PR 22-JUL-1999;	99US-0145087.	PR 21-OCT-1999;	99US-0160770.
PR 22-JUL-1999;	99US-0145089.	PR 21-OCT-1999;	99US-0160814.
PR 22-JUL-1999;	99US-0145192.	PR 21-OCT-1999;	99US-0160815.
PR 23-JUL-1999;	99US-0145145.	PR 22-OCT-1999;	99US-0160980.
PR 23-JUL-1999;	99US-0145218.	PR 22-OCT-1999;	99US-0160981.
PR 23-JUL-1999;	99US-0145224.	PR 22-OCT-1999;	99US-0160989.
PR 26-JUL-1999;	99US-0145276.	PR 25-OCT-1999;	99US-0161404.
PR 27-JUL-1999;	99US-0145913.	PR 25-OCT-1999;	99US-0161405.
PR 27-JUL-1999;	99US-0145918.	PR 25-OCT-1999;	99US-0161406.
PR 27-JUL-1999;	99US-0145919.	PR 26-OCT-1999;	99US-0161359.
PR 27-JUL-1999;	99US-0145951.	PR 26-OCT-1999;	99US-0161360.
PR 28-JUL-1999;	99US-0145951.	PR 26-OCT-1999;	99US-0161361.
PR 02-AUG-1999;	99US-0146386.	PR 28-OCT-1999;	99US-0161920.
PR 02-AUG-1999;	99US-0146388.	PR 28-OCT-1999;	99US-0161922.
PR 03-AUG-1999;	99US-0147038.	PR 28-OCT-1999;	99US-0161993.
PR 04-AUG-1999;	99US-0147204.	PR 29-OCT-1999;	99US-0161993.
PR 04-AUG-1999;	99US-0147302.	PR 29-OCT-1999;	99US-0162142.
PR 05-AUG-1999;	99US-0147192.		
PR 05-AUG-1999;	99US-0147260.		
PR 06-AUG-1999;	99US-0147303.		
PR 06-AUG-1999;	99US-0147416.		
PR 08-AUG-1999;	99US-0147493.		
PR 08-AUG-1999;	99US-0147935.		
PR 10-AUG-1999;	99US-0148171.		
PR 11-AUG-1999;	99US-0148319.		
PR 12-AUG-1999;	99US-0148341.		
PR 13-AUG-1999;	99US-0148565.		
PR 13-AUG-1999;	99US-0148684.		
PR 16-AUG-1999;	99US-0149368.		
PR 17-AUG-1999;	99US-0149175.		
PR 18-AUG-1999;	99US-0149426.		
PR 20-AUG-1999;	99US-0149722.		
PR 20-AUG-1999;	99US-0149723.		
PR 20-AUG-1999;	99US-0149929.		
PR 23-AUG-1999;	99US-0149902.		
PR 23-AUG-1999;	99US-0149930.		
PR 25-AUG-1999;	99US-0150566.		
PR 26-AUG-1999;	99US-0150884.		
PR 27-AUG-1999;	99US-0151065.		
PR 27-AUG-1999;	99US-0151066.		
PR 27-AUG-1999;	99US-0151080.		
PR 30-AUG-1999;	99US-0151303.		
PR 31-AUG-1999;	99US-0151438.		
PR 01-SEP-1999;	99US-0151920.		
PR 07-SEP-1999;	99US-0152363.		
PR 10-SEP-1999;	99US-0153070.		
PR 13-SEP-1999;	99US-0153758.		
PR 15-SEP-1999;	99US-0154018.		
PR 16-SEP-1999;	99US-0154039.		
PR 20-SEP-1999;	99US-0154779.		
PR 22-SEP-1999;	99US-0155139.		
PR 06-JUL-1999;	99US-0142390.	PR 23-SEP-1999;	99US-0155486.
PR 08-JUL-1999;	99US-0142803.	PR 24-SEP-1999;	99US-0155659.
PR 09-JUL-1999;	99US-0142920.	PR 28-SEP-1999;	99US-0156458.
PR 12-JUL-1999;	99US-0143277.	PR 29-SEP-1999;	99US-0156596.
PR 13-JUL-1999;	99US-0143542.	PR 04-OCT-1999;	99US-0157117.
PR 14-JUL-1999;	99US-0143624.	PR 05-OCT-1999;	99US-0157753.
PR 15-JUL-1999;	99US-0144005.	PR 06-OCT-1999;	99US-0157865.
PR 16-JUL-1999;	99US-0144085.	PR 07-OCT-1999;	99US-0158029.
PR 16-JUL-1999;	99US-0144086.	PR 08-OCT-1999;	99US-0158232.
PR 19-JUL-1999;	99US-0144325.	PR 12-OCT-1999;	99US-0158369.
PR 19-JUL-1999;	99US-0144331.	PR 13-OCT-1999;	99US-0159293.
PR 19-JUL-1999;	99US-0144332.	PR 13-OCT-1999;	99US-0159294.
PR 19-JUL-1999;	99US-0144333.	PR 13-OCT-1999;	99US-0159295.
PR 19-JUL-1999;	99US-0144334.	PR 14-OCT-1999;	99US-0159329.
PR 19-JUL-1999;	99US-0144335.	PR 14-OCT-1999;	99US-0159330.
PR 20-JUL-1999;	99US-0144352.	PR 14-OCT-1999;	99US-0159331.
PR 20-JUL-1999;	99US-0144632.	PR 14-OCT-1999;	99US-0159637.
PR 20-JUL-1999;	99US-0144884.	PR 14-OCT-1999;	99US-0159638.
PR 21-JUL-1999;	99US-0144814.	PR 18-OCT-1999;	99US-0159584.
PR 21-JUL-1999;	99US-0145086.	PR 21-OCT-1999;	99US-0160741.
PR 22-JUL-1999;	99US-0145088.	PR 21-OCT-1999;	99US-0160767.
PR 22-JUL-1999;	99US-0145085.	PR 21-OCT-1999;	99US-0160768.
PR 22-JUL-1999;	99US-0145087.	PR 21-OCT-1999;	99US-0160770.
PR 22-JUL-1999;	99US-0145089.	PR 21-OCT-1999;	99US-0160814.
PR 22-JUL-1999;	99US-0145192.	PR 21-OCT-1999;	99US-0160815.
PR 23-JUL-1999;	99US-0145145.	PR 22-OCT-1999;	99US-0160980.
PR 23-JUL-1999;	99US-0145218.	PR 22-OCT-1999;	99US-0160981.
PR 23-JUL-1999;	99US-0145224.	PR 22-OCT-1999;	99US-0160989.
PR 26-JUL-1999;	99US-0145276.	PR 25-OCT-1999;	99US-0161404.
PR 27-JUL-1999;	99US-0145913.	PR 25-OCT-1999;	99US-0161405.
PR 27-JUL-1999;	99US-0145918.	PR 25-OCT-1999;	99US-0161406.
PR 27-JUL-1999;	99US-0145919.	PR 26-OCT-1999;	99US-0161359.
PR 27-JUL-1999;	99US-0145951.	PR 26-OCT-1999;	99US-0161360.
PR 28-JUL-1999;	99US-0145951.	PR 26-OCT-1999;	99US-0161361.
PR 02-AUG-1999;	99US-0146386.	PR 28-OCT-1999;	99US-0161920.
PR 02-AUG-1999;	99US-0146388.	PR 28-OCT-1999;	99US-0161922.
PR 03-AUG-1999;	99US-0147038.	PR 28-OCT-1999;	99US-0161993.
PR 04-AUG-1999;	99US-0147204.	PR 29-OCT-1999;	99US-0161993.
PR 04-AUG-1999;	99US-0147302.		
PR 05-AUG-1999;	99US-0147192.		
PR 05-AUG-1999;	99US-0147260.		
PR 06-AUG-1999;	99US-0147303.		
PR 06-AUG-1999;	99US-0147416.		
PR 08-AUG-1999;	99US-0147493.		
PR 08-AUG-1999;	99US-0147935.		
PR 10-AUG-1999;	99US-0148171.		
PR 11-AUG-1999;	99US-0148319.		
PR 12-AUG-1999;	99US-0148341.		
PR 13-AUG-1999;	99US-0148565.		
PR 13-AUG-1999;	99US-0148684.		
PR 16-AUG-1999;	99US-0149368.		
PR 17-AUG-1999;	99US-0149175.		
PR 18-AUG-1999;	99US-0149426.		
PR 20-AUG-1999;	99US-0149722.		
PR 20-AUG-1999;	99US-0149723.		
PR 20-AUG-1999;	99US-0149929.		
PR 23-AUG-1999;	99US-0149902.		
PR 23-AUG-1999;	99US-0149930.		
PR 25-AUG-1999;	99US-0150566.		
PR 26-AUG-1999;	99US-0150884.		
PR 27-AUG-1999;	99US-0151065.		
PR 27-AUG-1999;	99US-0151066.		
PR 27-AUG-1999;	99US-0151080.		
PR 30-AUG-1999;	99US-0151303.		
PR 31-AUG-1999;	99US-0151438.		
PR 01-SEP-1999;	99US-0151920.		
PR 07-SEP-1999;	99US-0152363.		
PR 10-SEP-1999;	99US-0153070.		
PR 13-SEP-1999;	99US-0153758.		
PR 15-SEP-1999;	99US-0154018.		
PR 16-SEP-1999;	99US-0154039.		
PR 20-SEP-1999;	99US-0154779.		
PR 22-SEP-1999;	99US-0155139.		

Query Match 78.7%; Score 37; DB 21; Length 111;

Best Local Similarity 77.8%; Pred. No. 1.2e+02;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 STAEGGGGR 9

Db 82 SRAAGGGGR 90

RESULT 14

AAGS9596

ID AAGS9596 standard; Protein; 120 AA.

XX AC AAGS9596;

XX DT 18-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 77101.

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

Fri Dec 12 13:22:03 2003

```

PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 78.7%; Score 37; DB 21; Length 120;
Best Local Similarity 77.8%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 STAEAGGGR 9
Db 91 SAASGGGR 99

RESULT 15
AAW41611
ID AAW41611 standard; Protein; 536 AA.
XX AC AAW41611;
XX DT 20-APR-1998 (first entry)
XX DE Rape protox-1.
XX KW Protoporphyrinogen oxidase-1; protox-1; promoter; rape;
XX KW herbicide resistance; breeding programme; probe; gene isolation;
XX KW genomic mapping.
XX OS Brassica napus.
XX PN WO9732028-A1.
XX PD 04-SEP-1997.
XX PF 27-FEB-1997; 97WO-US03343.
XX PR 21-JUN-1996; 96US-0020003.
XX PR 28-FEB-1996; 96US-0012705.
XX PR 28-FEB-1996; 96US-0013612.
XX PA (NOVS ) NOVARTIS AG.
XX PI Johnson MA, Volrath SL, Ward ER;
XX DR WPI, 1997-489209/45.
XX DR N-PSDB; AAV04315.

```

```

XX DNA containing a plant proto-porphyrinogen oxidase gene promoter -
PT optionally linked to a heterologous gene, especially to express
PT herbicide-resistant enzymes, and plants containing such constructs
XX Claim 41; Pages 93-95; 114pp; English.
XX The present sequence is rape protoporphyrinogen oxidase-1
CC (protox-1).
CC The protox-1 promoter can be used to express herbicide resistant
CC enzymes, specifically protox, i.e. a plant tissue, plant or progeny
CC containing a chimeric gene of the promoter and a heterologous
CC coding sequence. The plant can also be used in breeding programmes.
CC Also hybridising fragments of the protox coding sequence can be
CC used as probes, e.g. to isolate related genes or for genomic
CC mapping.
XX Sequence 536 AA;
SQ
Query Match 78.7%; Score 37; DB 18; Length 536;
Best Local Similarity 87.5%; Pred. No. 5.1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 STAEAGGGR 8
Db 43 STIEGGGG 50

Search completed: December 11, 2003, 18:25:00
Job time : 27 secs

```

Query Match 85.1%; Score 40; DB 3; Length 95;
Best Local Similarity 88.9%; Pred. No. 8.4;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 STAEAGGGR 9
DB 21 SHAEGGGGR 29

RESULT 2
US-09-252-991A-23062
; Sequence 23062, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23062
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23062

Query Match 80.9%; Score 38; DB 4; Length 212;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AEGGGGR 9
DB 171 AEGGGGR 177

RESULT 3
US-09-252-991A-24374
; Sequence 24374, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24374
; LENGTH: 1116
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24374

Query Match 80.9%; Score 38; DB 4; Length 1116;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AEGGGGR 9
DB 734 AEGGGGR 740

RESULT 4
US-08-808-931-20

; Sequence 20, Application US/08808931
; Patent No. 5939602
; GENERAL INFORMATION:
; APPLICANT: Volrath, Sandra
; APPLICANT: Johnson, Marie
; APPLICANT: Potter, Sharon
; APPLICANT: Ward, Eric
; APPLICANT: Helfetz, Peter
; TITLE OF INVENTION: DNA Molecules Encoding Plant
; TITLE OF INVENTION: Protoporphyrinogen Oxidase and Inhibitor-Resistant Mutants
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: No. 5939602artis Corporation 2005
; STREET: 520 White Plains Road, P.O. Box 2005
; CITY: Tarrytown
; STATE: NY
; COUNTRY: USA
; ZIP: 10591-9005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,931
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/012,705
; FILING DATE: 28-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/013,612
; FILING DATE: 28-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/020,003
; FILING DATE: 21-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC 1847
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919) 541-8587
; TELEFAX: (919) 541-8689
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 536 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
US-08-808-931-20

Query Match 78.7%; Score 37; DB 2; Length 536;
Best Local Similarity 87.5%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 STAEAGGGR 8
DB 43 STIEGGGG 50

RESULT 5
US-08-808-323-20
; Sequence 20, Application US/08808323
; Patent No. 6018105
; GENERAL INFORMATION:
; APPLICANT: Johnson, Marie
; APPLICANT: Volrath, Sandra
; APPLICANT: Ward, Eric
; TITLE OF INVENTION: Promoters from Plant
; TITLE OF INVENTION: Protoporphyrinogen Oxidase Genes
; NUMBER OF SEQUENCES: 26

```

; FILING DATE: 13-APR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/050,603
; FILING DATE: 30-MAR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/126,430
; FILING DATE: 11-MAR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/808,931
; FILING DATE: 28-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/012,705
; FILING DATE: 28-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/013,612
; FILING DATE: 28-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/020,003
; FILING DATE: 21-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/472,028
; FILING DATE: 06-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC 1847/CIP3
; TELEPHONE: (919) 541-8587
; TELEFAX: (919) 541-8689
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 536 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; US-09-102-420B-20

Query Match 78.7%; Score 37; DB 3; Length 536;
Best Local Similarity 87.5%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 STAEGGG 8
DB 43 STIEGGG 50

RESULT 8
US-09-497-698-20
; Sequence 20, Application US/09497698
; Patent No. 6308458
; GENERAL INFORMATION:
; APPLICANT: Volrath, Sandra
; Johnson, Marie
; Ward, Eric
; Heifetz, Peter
; TITLE OF INVENTION: HERBICIDE-TOLERANT PROTOPORPHYRINOGEN
; OXIDASE ("PROTOX")
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6308458artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/497,698

```

```

; FILING DATE: 03-FEB-2000
; CLASSIFICATION: <Unknown>
; 30-MAR-1998
; 11-MAR-1998
; 28-FEB-1997
; 28-FEB-1996
; 28-FEB-1996
; 21-JUN-1996
; 06-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/102,420
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 09/050,603
; FILING DATE: 30-MAR-1998
; APPLICATION NUMBER: US 60/126,430
; FILING DATE: 11-MAR-1998
; APPLICATION NUMBER: US 08/808,931
; FILING DATE: 28-FEB-1997
; APPLICATION NUMBER: US 60/012,705
; FILING DATE: 28-FEB-1996
; APPLICATION NUMBER: US 60/013,612
; FILING DATE: 28-FEB-1996
; APPLICATION NUMBER: US 60/020,003
; FILING DATE: 21-JUN-1996
; APPLICATION NUMBER: US 08/472,028
; FILING DATE: 06-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC 1847/CIP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919) 541-8587
; TELEFAX: (919) 541-8689
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 536 amino acids
; TYPE: amino acid
; STRANDEDNESS: No. 6308458 Relevant
; TOPOLOGY: No. 6308458 Relevant
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-497-698-20

Query Match 78.7%; Score 37; DB 4; Length 536;
Best Local Similarity 87.5%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 STAEGGG 8
DB 43 STIEGGG 50

RESULT 9
US-08-406-824A-17
; Sequence 17, Application US/08406824A
; Patent No. 6541810
; GENERAL INFORMATION:
; APPLICANT: SMITH, Craig A.
; TITLE OF INVENTION: TUMOR NECROSIS FACTOR-ALPHA AND BETA-RECEPTORS
; FILE REFERENCE: A-71592
; CURRENT APPLICATION NUMBER: US/08/406,824A
; CURRENT FILING DATE: 1995-03-20
; PRIOR APPLICATION NUMBER: US 08/255,849
; PRIOR FILING DATE: 1994-06-08
; PRIOR APPLICATION NUMBER: US 07/860,710
; PRIOR FILING DATE: 1992-03-30
; PRIOR APPLICATION NUMBER: US 07/523,635
; PRIOR FILING DATE: 1990-05-10
; PRIOR APPLICATION NUMBER: US 07/421,417
; PRIOR FILING DATE: 1989-10-13
; PRIOR APPLICATION NUMBER: US 07/405,370
; PRIOR FILING DATE: 1989-09-11
; PRIOR APPLICATION NUMBER: US 07/403,241

```

GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25950
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25950

Query Match 72.3%; Score 34; DB 4; Length 144;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 EGGGR 9
Db 28 EGGGR 33

RESULT 14
US-09-252-991A-17577
; Sequence 17577, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17577
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17577

Query Match 72.3%; Score 34; DB 4; Length 161;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 EGGGR 9
Db 114 EGGGR 119

RESULT 15
US-08-906-769-145
; Sequence 145, Application US/08906769
; Patent No. 6077687
; GENERAL INFORMATION:
; APPLICANT: Grieve, Robert B.
; APPLICANT: Rushlow, Keith E.
; APPLICANT: Wu Hunter, Shirley
; APPLICANT: Frank, Glenn R.
; APPLICANT: Sciegler, Gary
; APPLICANT: Gaines, Patrick J.
; TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
; MOLECULES AND USES THEREOF
; NUMBER OF SEQUENCES: 190

CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patenlin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/906,769
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/639,075
; FILING DATE: 24-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-25-C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 145:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 267 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-906-769-145

Query Match 72.3%; Score 34; DB 3; Length 267;
Best Local Similarity 75.0%; Pred. No. 2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 STAEGGGG 8
Db 161 ATKEGGGG 168

Search completed: December 11, 2003, 18:30:41
Job time : 10.2222 secs

```

; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; PRIOR FILING DATE: 2001-08-10
; CURRENT APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 1200
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (16)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1200

```

```

Query Match      83.0%; Score 39; DB 10; Length 174;
Best Local Similarity 87.5%; Pred. No. 85;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 2 TAEAGGGGR 9
| | | | |
Db 7 TTEGGGGR 14

```

```

RESULT 3
US-10-029-386-33773
; Sequence 33773, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David K.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AROMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33773
; LENGTH: 76
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO ACC06050.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.74
US-10-029-386-33773

```

```

Query Match      90.9%; Score 38; DB 12; Length 76;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 3 AEGGGGR 9
| | | | |
Db 61 AEGGGGR 67

```

```

RESULT 4
US-10-259-165-10
; Sequence 10, Application US/10259165
; Publication No. US2003013588A1
; GENERAL INFORMATION:
; APPLICANT: Zhu, Tong
; APPLICANT: Wang, Xun
; APPLICANT: Chang, Hur-song
; APPLICANT: Briggs, Steven P.

```

```

; APPLICANT: Cooper, Bret
; APPLICANT: Glazebrook, Jane
; APPLICANT: Goff, Stephen A.
; APPLICANT: Katagiri, Fumiaki
; APPLICANT: Kreps, Joel
; APPLICANT: Moughamer, Todd
; APPLICANT: Provart, Nicholas
; APPLICANT: Ricke, Darrell
; TITLE OF INVENTION: GENES THAT ARE MODULATED BY POSTTRANSCRIPTIONAL GENE SILENCING
; FILE REFERENCE: 70030-NP
; CURRENT APPLICATION NUMBER: US/10/259,165
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 60/368,327
; PRIOR FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; NUMBER OF SEQ ID NOS: 782
; SOFTWARE: PatentList.pl version 3.0.4 (C) 2001 Syngenta
; SEQ ID NO 10
; LENGTH: 157
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: X region
; LOCATION: (33)..(34)
; OTHER INFORMATION: Xaa = any naturally occurring amino acid
; NAME/KEY: X region
; LOCATION: (36)..(36)
; OTHER INFORMATION: Xaa = any naturally occurring amino acid
US-10-259-165-10

```

```

Query Match      80.9%; Score 38; DB 12; Length 157;
Best Local Similarity 100.0%; Pred. No. 11e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 3 AEGGGGR 9
| | | | |
Db 107 AEGGGGR 113

```

```

RESULT 5
US-10-259-165-344
; Sequence 344, Application US/10259165
; Publication No. US2003013588A1
; GENERAL INFORMATION:
; APPLICANT: Zhu, Tong
; APPLICANT: Wang, Xun
; APPLICANT: Chang, Hur-song
; APPLICANT: Briggs, Steven P.
; APPLICANT: Cooper, Bret
; APPLICANT: Glazebrook, Jane
; APPLICANT: Goff, Stephen A.
; APPLICANT: Katagiri, Fumiaki
; APPLICANT: Kreps, Joel
; APPLICANT: Moughamer, Todd
; APPLICANT: Provart, Nicholas
; APPLICANT: Ricke, Darrell
; TITLE OF INVENTION: GENES THAT ARE MODULATED BY POSTTRANSCRIPTIONAL GENE SILENCING
; FILE REFERENCE: 70030-NP
; CURRENT APPLICATION NUMBER: US/10/259,165
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 60/368,327
; PRIOR FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; NUMBER OF SEQ ID NOS: 782
; SOFTWARE: PatentList.pl version 3.0.4 (C) 2001 Syngenta
; SEQ ID NO 344

```

Fri Dec 12 13:22:06 2003

ADDRESSEE: No. US20020073443Alartis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: NC
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/730,917
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/059,164
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/038,878
FILING DATE: 11-MAR-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/808,931
FILING DATE: 28-FEB-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/012,705
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/013,612
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/020,003
FILING DATE: 21-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: CGC1847/CIP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 536 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
US-09-730-917-20

Query Match 78.7%; Score 37; DB 9; Length 536;
Best Local Similarity 87.5%; Pred. No. 4.7e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 STAEGGGG 8
Db 43 STIEGGGG 50

RESULT 9
US-10-000-256A-173
; Sequence 173, Application US/10000256A
; Publication No. US20030039983A1
; GENERAL INFORMATION:
; APPLICANT: Sun, Yongming
; APPLICANT: Recipon, Hervé
; APPLICANT: Chen, Sei-yu
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific Genes and
; FILE REFERENCE: DEX-0259
; CURRENT APPLICATION NUMBER: US/10/000,256A
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: 60/244,782
; PRIOR FILING DATE: 2000-11-01

NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 173
; LENGTH: 121
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-000-256A-173

Query Match 76.6%; Score 36; DB 15; Length 121;
Best Local Similarity 66.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 STAEGGGG 9
Db 69 SSGQGGGG 77

RESULT 10
US-10-021-811-26
; Sequence 26, Application US/10021811
; Publication No. US20030024007A1
; GENERAL INFORMATION:
; APPLICANT: Caboon, Rebecca E.
; APPLICANT: Fang, Yiwen
; APPLICANT: Odell, Joan
; APPLICANT: Weng, Zude
; TITLE OF INVENTION: Plant Myb Transcription Factor Homologs
; FILE REFERENCE: BB1294 US NA
; CURRENT APPLICATION NUMBER: US/10/021,811
; CURRENT FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: 60/110,609
; PRIOR FILING DATE: 1998-December-02
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 26
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-021-811-26

Query Match 76.6%; Score 36; DB 15; Length 323;
Best Local Similarity 66.7%; Pred. No. 4.2e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 STAEGGGG 9
Db 130 AASEGGGG 138

RESULT 11
US-10-156-761-12539
; Sequence 12539, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 243-282
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 12539
; LENGTH: 726
; TYPE: PRT

Fri Dec 12 13:22:06 2003

```
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 522
; LENGTH: 156
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-10-101-464A-522
```

```
Query Match 74.5%; Score 35; DB 15; Length 156;
Best Local Similarity 85.7%; Pred. No. 3.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3 AEGGGGR 9
Db 133 ADGGGGR 139
```

```
RESULT 15
US-10-259-165-216
; Sequence 216, Application US/10259165
; Publication No. US20030135888A1
; GENERAL INFORMATION:
; APPLICANT: Zhu, Tong
; APPLICANT: Wang, Xun
; APPLICANT: Chang, Hur-song
; APPLICANT: Briggs, Steven P.
; APPLICANT: Cooper, Bret
; APPLICANT: Glazebrook, Jane
; APPLICANT: Goff, Stephen A.
; APPLICANT: Katagiri, Fumiyaki
; APPLICANT: Kreps, Joel Todd
; APPLICANT: Moughamer,
; APPLICANT: Provart, Nicholas
; APPLICANT: Rickel, Darrell
; TITLE OF INVENTION: GENES THAT ARE MODULATED BY POSTTRANSCRIPTIONAL GENE SILENCING
; FILE REFERENCE: 70030-NP
; CURRENT APPLICATION NUMBER: US/10/259,165
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 60/368,327
; PRIOR FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; NUMBER OF SEQ ID NOS: 782
; SOFTWARE: PatentList.pl version 3.0.4 (C) 2001 Syngenta
; SEQ ID NO 216
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-259-165-216
```

```
Query Match 74.5%; Score 35; DB 12; Length 352;
Best Local Similarity 85.7%; Pred. No. 5.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3 AEGGGGR 9
Db 69 ADGGGGR 75
```

Search completed: December 11, 2003, 18:38:04
Job time : 49.3333 secs

Query Match 80.9%; Score 38; DB 2; Length 1078;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AEGGGGR 9
 |||||
 Db 699 AEGGGGR 705

RESULT 3
 C75585
 carboxylase - Deinococcus radiodurans (strain R1)
 C:Species: Deinococcus radiodurans
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
 C:Accession: C75585
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
 M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A:Reference number: A75250; MUID:20036896; PMID:10567266
 A:Accession: C75585
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1091 <WHI>
 A:Cross-references: GB:AE001863; GB:AE001825; NID:G6460670; PIDN:AAF12468.1; PID:G646076
 A:Experimental source: strain R1
 C:Genetics:
 A:Gene: DRA0310
 A:Map position: 2

Query Match 80.9%; Score 38; DB 2; Length 1091;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AEGGGGR 9
 |||||
 Db 711 AEGGGGR 717

RESULT 4
 B83471
 Probable pyruvate carboxylase PA1400 [imported] - Pseudomonas aeruginosa (strain PA01)
 C:Species: Pseudomonas aeruginosa
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C:Accession: B83471
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
 ; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
 A:Reference number: A82950; MUID:20437337; PMID:10984043
 A:Accession: B83471
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1095 <STO>
 A:Cross-references: GB:AE004569; GB:AE004091; NID:G9947339; PIDN:AAG04789.1; GSPDB:GN001
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: PA1400

Query Match 80.9%; Score 38; DB 2; Length 1095;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AEGGGGR 9
 |||||
 Db 713 AEGGGGR 719

RESULT 5
 T49330
 cytochrome inhibitor byr4 related protein [imported] - Neurospora crassa

N:Alternate names: protein B13N20.150
 C:Species: Neurospora crassa
 C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
 C:Accession: T49330
 R:Schulte, U.; Aign, V.; Hchisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura
 submitted to the Protein Sequence Database, May 2000
 A:Reference number: Z25022
 A:Accession: T49330
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1172 <SCH>
 A:Cross-references: EMBL:ALJ55925; GSPDB:GN00116; NCSP:B13N20.150
 A:Experimental source: BAC clone B13N20; strain OR74A
 C:Genetics:
 A:Gene: NCSP:B13N20.150
 A:Map position: 6
 A:Introns: 862/1; 1038/1; 1062/2

Query Match 80.9%; Score 38; DB 2; Length 1172;
 Best Local Similarity 77.8%; Pred. No. 1.5e+02;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 STAEGGGGR 9
 |||||
 Db 1068 STNGGGGR 1076

RESULT 6
 C71401
 Hypothetical protein - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cross)
 A:Variety: Columbia
 C:Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 05-Dec-1998
 C:Accession: C71401
 R:Byan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dir
 P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenecker, T.; Pohl, T.M.; Terry, N.; Gie
 avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
 Nature 391, 485-488, 1998
 A:Authors: Mueller-Auer, S.; Silvey, M.; Jones, R.; Montfort, A.; Pons, A.; Puigdomenech
 erhoit, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; An
 C.; Chalwatzis, N.
 A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis tha
 A:Reference number: A71400; MUID:98121113; PMID:9461215
 A:Accession: C71401
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-111 <BEV>
 A:Cross-references: GB:Z97335; NID:G2244747; PID:e326863; PID:G2244758
 C:Genetics:
 A:Map position: 4COP9-4G3845

Query Match 78.7%; Score 37; DB 2; Length 111;
 Best Local Similarity 77.8%; Pred. No. 24;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 STAEGGGGR 9
 |||||
 Db 82 SAASGGGR 90

RESULT 7
 E87275
 TonB-dependent receptor [imported] - Caulobacter crescentus
 C:Species: Caulobacter crescentus
 C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
 C:Accession: E87275
 R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolo
 n, J.; Ermolaeva, N.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M
 Proc. Natl. Acad. Sci. U.S.A. 98, 4138-4141, 2001
 A:Title: Complete Genome Sequence of Caulobacter crescentus.
 A:Reference number: A87249; MUID:21173698; PMID:11259647
 A:Accession: E87275

C:Genetics:
 A:Introns: 4/3
 C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain
 C:Keywords: hydrolase

Query Match 74.5%; Score 35; DB 2; Length 39;

Best Local Similarity 66.7%; Pred. No. 20;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 STAEGGGGR 9

Db 14 TTAENGKK 22

RESULT 13

E72459

hypothetical protein APE2321 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix

C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999

C:Accession: E72459

R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K

DNA Res. 6, 83-101, 1999

A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr

A;Reference number: A72450; MUID:99310339; PMID:10382966

A;Accession: E72459

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-183 <RAW>

A;Cross-references: DDBJ:AP000064; NID:G5105945; PIDN:BAA81333.1; PID:d1045119; PID:G510

A;Experimental source: strain K1

C:Genetics:

A;Gene: APE2321

Query Match

Best Local Similarity 74.5%; Score 35; DB 2; Length 183;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 STAEGGGGR 9

Db 76 STAPGGGR 84

RESULT 14

E48423

homeotic protein engrailed 2 - human

N;Alternate names: homeotic protein En-2

C:Species: Homo sapiens (man)

C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 17-Oct-1997

C:Accession: E48423; A30141

R;Logan, C.; Hanks, M.C.; Noble-Topham, S.; Nailainathan, D.; Provart, N.J.; Joyner, A.I

Dev. Genet. 13, 345-358, 1992

A;Title: Cloning and sequence comparison of the mouse, human, and chicken engrailed gene

A;Reference number: A48423; MUID:93185339; PMID:1363401

A;Accession: E48423

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: DNA

A;Residues: 1-333 <LOG>

R;Poole, S.J.; Law, M.L.; Kao, F.T.; Lau, Y.F.

Genomics 4, 225-231, 1989

A;Title: Isolation and chromosomal localization of the human En-2 gene.

A;Reference number: A30141; MUID:89233109; PMID:2565873

A;Accession: A30141

A;Molecule type: DNA

A;Residues: 230-333 <POO>

C:Genetics:

A;Gene: En-2

C:Superfamily: unassigned homeobox proteins; homeobox homology

C:Keywords: DNA binding; homeobox; nucleus; transcription regulation

F;245-301/domain: homeobox homology <HOX>

Query Match

Best Local Similarity 74.5%; Score 35; DB 2; Length 333;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 STAEGGGGR 8

Db 107 SGREGGGG 114

RESULT 15

T29121

hypothetical protein SC1F2.12 - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999

C:Accession: T29121

R;Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, August 1998

A;Reference number: Z17215

A;Accession: T29121

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-400 <PAR>

A;Cross-references: EMBL:AL031350; NID:el316892; PID:el316904; PIDN:CAA20503.1

C:Genetics:

A;Note: SC1F2.12

Query Match

Best Local Similarity 74.5%; Score 35; DB 2; Length 400;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TAEGGGGR 9

Db 198 TAEGGGGR 205

Search completed: December 11, 2003, 18:29:42

Job time : 11.3333 secs

```

CC -----
DR EMBL; Z74410; CAA98927.1; -
DR EMBL; AB006921; AAK44323.1; -
DR PIR; C70750; C70750.
DR TIGR; WTC100; -
DR Tuberculist; Rv0091; -
DR InterPro; IPR000845; PNP_UDP.
DR Pfam; PF01048; PNP_UDP.1; 1.
DR ProDom; PD003928; PNP_UDP; 1.
KW Hydrolase; Multifunctional enzyme; Complete proteome.
SQ SEQUENCE 255 AA; 27339 MW; EA80A8CB2BFEE79 CRC64;

Query Match      87.2%; Score 41; DB 1; Length 255;
Best Local Similarity 88.9%; Pred. No. 4.3;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 STAEQGGGR 9
   ||| |||||
DB 152 STAEQGGGR 160

RESULT 2
RELB_CHK
ID RELB_CHK     STANDARD;      PRT; 549 AA.
AC P51509; Q90721; Q9PWF4;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Transcription factor RelB homolog.
GN RELB.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OC NCBI_TaxID=9031;
OX [1]
RN SEQUENCE FROM N.A.
RP TISSUE=Spleen;
RC Piffat K.A., Ikeda T., Hrdlickova R., Nehyba J., Liss A., Huang S.,
RA Sif S., Gilmore T.D., Bose H.R.;
RT "Characterization of the chicken RelB transcription factor.";
RN Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE OF 151-424 FROM N.A.
RP TISSUE=Spleen;
RC Ikeda T.;
RA Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
[3]
RN SEQUENCE OF 150-479 FROM N.A.
RA Gilmore T.D., Piffat K.A., Huang S., Sif S.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: STIMULATES PROMOTER ACTIVITY IN THE PRESENCE OF P49- AND
CC P50-NF-KAPPA-B. NEITHER ASSOCIATES WITH DNA NOR WITH P65-NF-KAPPA-
CC B (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: BELONGS TO THE REL/DORSAL FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF029260; AAD41539.1; -
DR EMBL; D13794; BAA02947.1; -
DR HSP; P25799; IBFT.
DR InterPro; IPR002909; IPT_TIG.
DR InterPro; IPR000451; NF_Rel_dor_fam.
DR Pfam; PF00554; RHD; 1.
DR Pfam; PF01833; TIG; 1.

```

```

DR PRINTS; PR00057; NFKB1NSCPFCT.
DR SMART; SMO0429; IPT; 1.
DR PROSITE; PS01284; REL_1; 1.
DR PROSITE; PS0254; REL_2; 1.
KW Nuclear protein; Transcription regulation; Activator.
FT DOMAIN 135 412 REL-Like (RHD).
FT CONFLICT 266 266 G -> A (IN REF. 1).
FT CONFLICT 364 364 A -> G (IN REF. 2).
FT CONFLICT 370 370 G -> E (IN REF. 2).
FT CONFLICT 380 380 R -> A (IN REF. 3).
FT CONFLICT 479 479 S -> F (IN REF. 3).
SQ SEQUENCE 549 AA; 17668CF2C78BA37D CRC64;

Query Match      85.1%; Score 40; DB 1; Length 549;
Best Local Similarity 87.5%; Pred. No. 14;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 TAEQGGGR 9
   ||| |||||
DB 520 TAEQGGGR 527

RESULT 3
HEM6_XANCP
ID HEM6_XANCP     STANDARD;      PRT; 299 AA.
AC Q8P3Q0;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Coproporphyrinogen III oxidase, aerobic [EC 1.3.3.3]
DE (Coproporphyrinogenase) (Coprogen oxidase).
GN HEMF OR XCC4019.
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OC NCBI_TaxID=340;
OX [1]
RN SEQUENCE FROM N.A.
RP STRAIN=ATCC 33913 / NCPPB 528;
RC MEDLINE=20022145; PubMed=12024217;
RA da Silva A.C.B., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.C., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavan F., Cardozo J., Chabergo F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Gruber A.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Lemos M.V.F.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spindola L.A.F., Takita M.A., Tamura R.B., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities.";
RL Nature 417:459-463(2002).
CC -1- CATALYTIC ACTIVITY: Coproporphyrinogen-III + O(2) =
CC protoporphyrinogen-IX + 2 CO(2).
CC -1- COFACTOR: Iron (By similarity).
CC -1- PATHWAY: Porphyrin biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE AEROBIC COPROPORPHYRINOGEN III OXIDASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

Query Match 76.6%; Score 36; DB 1; Length 410;
 Best Local Similarity 100.0%; Pred. No. 47;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STAE6GG 7
 |||||
 DB 175 STAE6GG 181

RESULT 6

SRF_DROME STANDARD; PRT; 450 AA.
 AC Q24535;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Serum response factor homolog (dsrf) (Blasted protein).
 GN SRF OR SRF.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]

SEQUENCE FROM N.A. AND CHARACTERIZATION.

RP MEDLINE=95324363; PubMed=7600954;
 RX Afoelter M., Montagne J., Walldorf U., Gropp J.C., Kloter U.,
 RA Larosa M., Gehring W.J.;
 RT "The Drosophila SRF homolog is expressed in a subset of tracheal
 cells and maps within a genomic region required for tracheal
 development";
 RL Development 120:743-753(1994).

CC -1- FUNCTION: MIGHT PLAY A ROLE IN THE PROPER FORMATION AND

CC MAINTENANCE OF THE TRACHEA.

CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).

CC -1- DEVELOPMENTAL STAGE: EXPRESSED DURING SEVERAL PHASES OF EMBRYONIC
 CC DEVELOPMENT. IN THE EGG, BOTH THE RNA AND THE PROTEIN ARE MATERNAL
 CC IN ORIGIN AND SLOWLY DECREASE IN AMOUNT DURING GASTRULATION. AFTER
 CC GERM BAND RETRACTION, HIGH LEVELS OF ZYGOTIC EXPRESSION ARE
 CC OBSERVED IN A DISTINCT SUBSET OF PERIPHERAL TRACHEAL CELLS
 CC DISTRIBUTED THROUGHOUT THE EMBRYO AND LOW LEVELS IN SOMATIC
 CC MUSCLE.

CC -1- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION

CC FACTORS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; X77532; CAA54670.1; -;
 CC PIR; S42825; S42825.
 CC HSP; FL1831; 1SR5.
 CC TRANSFAC; T03688; -;
 CC FlyBase; FBgn0004101; bs.
 CC GO; GO:0005634; C:nucleus; IDA.
 CC GO; GO:0007430; P:terminal branching of trachea, cytoplasmic . . .; IMP.
 CC InterPro; IPR002100; TF_MADSbox.
 CC Pfam; PF00319; SRF-TF; 1
 CC PRINTS; PR00404; MADSDOMAIN.
 CC SMART; SM00432; MADS; 1.

CC PROSITE; PS00350; MADS_BOX_1; 1.

CC PROSITE; PS00066; MADS_BOX_2; 1.

CC KW Transcription regulation; DNA-binding; Activator; Nuclear protein;
 CC Developmental protein.

FT DOMAIN 167 221 MADS

FT DOMAIN 81 102 POLY-GLN.

FT DOMAIN 357 360 POLY-GLY.

SQ SEQUENCE 450 AA; 47777 MW; F49BF85ED597D3AE CRC64;

Query Match 76.6%; Score 36; DB 1; Length 450;
 Best Local Similarity 87.5%; Pred. No. 52;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 STAE6GG 8
 |||||
 DB 353 STAE6GG 360

RESULT 7

CPSA_MOUSE STANDARD; PRT; 1441 AA.
 AC Q9EP04;
 DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Cleavage and polyadenylation specificity factor, 160 kDa subunit (CPSF

DE 160 kDa subunit).

GN CPSF1 OR CPSF160.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Testis;

RA MEDLINE=21261509; PubMed=11369601;

RT "Overexpression of the Cstf-64 and CPSF-160 polyadenylation protein

RL Biol. Reprod. 64:1722-1729(2001).

CC -1- FUNCTION: CPSF PLAYS A KEY ROLE IN PRE-MRNA 3'-END FORMATION,

CC RECOGNIZING THE AAUAAA SIGNAL SEQUENCE AND INTERACTING WITH

CC POLY(A) POLYMERASE AND OTHER FACTORS TO BRING ABOUT CLEAVAGE AND

CC POLY(A) ADDITION. THIS SUBUNIT IS INVOLVED IN THE RNA RECOGNITION

CC STEP OF THE POLYADENYLATION REACTION (By similarity).

CC -1- SUBUNIT: CPSF IS A HETEROTETRAMER COMPOSED OF FOUR DISTINCT

CC SUBUNITS 160, 100, 70 AND 30 kDa.

CC -1- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOPLASM (By similarity).

CC -1- SIMILARITY: BELONGS TO THE CPSF160 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; AF322193; AAG40326.1; -;

CC InterPro; IPR004871; CPSF_A.

CC Pfam; PF03178; CPSF_A; 1.

CC KW mRNA processing; Nuclear protein; RNA-binding.

SQ SEQUENCE 1441 AA; 160817 MW; 5D927224152AC3B9 CRC64;

Query Match 76.6%; Score 36; DB 1; Length 1441;

Best Local Similarity 77.8%; Pred. No. 1.7e+02;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 STAE6GG 9

DB 915 STE6SGGR 923

RESULT 8

HME2_HUMAN STANDARD; PRT; 333 AA.

ID P19622;

DT 01-FEB-1991 (Rel. 17, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DE Homeobox protein engrailed-2 (Hu-En-2).

GN EN2.

CC II SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL: AJ223599; CAAL1465.1; -

CC DR HSP; P04191; IEUL.

CC DR InterPro: IPR001757; ATPase_E1-E2.

CC DR InterPro: IPR006069; Cation_ATPase.

CC DR InterPro: IPR006068; Cation_ATPase_C.

CC DR InterPro: IPR004014; Cation_ATPase_N.

CC DR InterPro: IPR005834; Hydrolase.

CC DR InterPro: IPR005775; Na/K ATPase alph.

CC DR Pfam: PF00689; Cation_ATPase_C; 1.

CC DR Pfam: PF00690; Cation_ATPase_N; 1.

CC DR Pfam: PF00122; E1-E2_ATPase; 1.

CC DR Pfam: PF00702; Hydrolase; 1.

CC DR PRINTS: PR00119; CATATPASE.

CC DR PRINTS: PR00121; NAKATPASE.

CC DR TIGRfams: TIGR01106; ATPase-IIC X-X; 1.

CC DR TIGRfams: TIGR01494; ATPase_P-type; 5.

CC DR PROSITE: PS00154; ATPase_E1-E2; 1.

CC KW Hydrolase; Sodium/potassium transport; Transmembrane; Phosphorylation;

CC KW Magnesium; Metal-binding; ATP-binding; Multigene family.

CC FT DOMAIN 1 82 CYTOPLASMIC (POTENTIAL).

CC FT TRANSMEM 83 103 POTENTIAL.

CC FT DOMAIN 104 126 LUMENAL (POTENTIAL).

CC FT TRANSMEM 127 147 POTENTIAL.

CC FT DOMAIN 148 283 CYTOPLASMIC (POTENTIAL).

CC FT TRANSMEM 284 303 POTENTIAL.

CC FT DOMAIN 304 315 LUMENAL (POTENTIAL).

CC FT TRANSMEM 316 333 POTENTIAL.

CC FT DOMAIN 334 766 CYTOPLASMIC (POTENTIAL).

CC FT TRANSMEM 767 786 POTENTIAL.

CC FT DOMAIN 787 796 LUMENAL (POTENTIAL).

CC FT TRANSMEM 797 817 POTENTIAL.

CC FT DOMAIN 818 837 CYTOPLASMIC (POTENTIAL).

CC FT TRANSMEM 838 860 POTENTIAL.

CC FT DOMAIN 861 912 LUMENAL (POTENTIAL).

CC FT TRANSMEM 913 932 POTENTIAL.

CC FT DOMAIN 933 945 CYTOPLASMIC (POTENTIAL).

CC FT TRANSMEM 946 964 POTENTIAL.

CC FT DOMAIN 965 979 LUMENAL (POTENTIAL).

CC FT TRANSMEM 980 1000 POTENTIAL.

CC FT DOMAIN 1001 1017 CYTOPLASMIC (POTENTIAL).

CC FT MOD RES 371 371 PHOSPHORYLATION (BY PKA) (BY SIMILARITY).

CC FT MOD RES 937 937 PHOSPHORYLATION (BY PKA) (BY SIMILARITY).

CC FT BINDING 77 79 BINDING OF PHOSPHOINOSITIDE-3 KINASE (BY SIMILARITY).

CC FT METAL 711 711 MAGNESIUM (BY SIMILARITY).

CC FT METAL 715 715 MAGNESIUM (BY SIMILARITY).

CC FT BINDING 502 502 ATP (BY SIMILARITY).

CC SQ SEQUENCE 1017 AA; 112050 MW; 3887CLBDE9386C5 CRC64;

Query Match 74.5%; Score 35; DB 1; Length 1017;

Best Local Similarity 66.7%; Pred. No. 1.7e+02;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 STAGGGGG 8

Db 268 STAGGGGG 275

RESULT 11

AL12 CHICK STANDARD; PRT; 1017 AA.

AC P24737;

DT 01-MAR-1992 (Rel. 21, Created)

DT 01-MAR-1992 (Rel. 21, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Sodium/potassium-transporting ATPase alpha-2 chain (EC 3.6.3.9)

DE (Sodium pump 2) (Na+/K+ ATPase 2).

GN ATP1A2.

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Gallus.

OC NCBI_TaxID=9031;

RN [1]

RP SEQUENCE FROM N.A.

RP MEDLINE=91023019; PubMed=2171348;

RA Takeyasu K., Lenas V., Fambrough D.M.;

RT "Stability of Na(+)-K(+)-ATPase alpha-subunit isoforms in evolution.";

RL Am. J. Physiol. 259:C619-C630(1990).

CC -!- FUNCTION: THIS IS THE CATALYTIC COMPONENT OF THE ACTIVE ENZYME, OF WHICH CATALYZES THE HYDROLYSIS OF ATP COUPLED WITH THE EXCHANGE OF NA AND K IONS ACROSS THE PLASMA MEMBRANE. THIS ACTION CREATES THE ELECTROCHEMICAL GRADIENT OF NA AND K, PROVIDING THE ENERGY FOR ACTIVE TRANSPORT OF VARIOUS NUTRIENTS.

CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + Na(+) (In) + K(+) (Out) = ADP + phosphate + Na(+) (Out) + K(+) (In).

CC -!- SUBUNIT: COMPOSED OF THREE SUBUNITS: ALPHA (CATALYTIC), BETA AND GAMMA.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein.

CC -!- SIMILARITY: Belongs to the cation transport ATPases family (P-type ATPases). Subfamily IIC.

ID A1A2 RAT STANDARD; PRT; 1020 AA.
AC P06686;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DE Sodium/potassium-transporting ATPase alpha-2 chain precursor
DE (EC 3.6.3.9) (Sodium pump 2) (Na+/K+ ATPase 2) (Alpha(+)).
GN ATP1A2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87128908; PubMed=3028470;
RA Shull G.E., Greb J., Lingrel J.B.;
RT "Molecular cloning of three distinct forms of the Na+,K+-ATPase
RT alpha-subunit from rat brain.";
RL Biochemistry 25:8125-8132(1986).
RN [2]
RP SEQUENCE OF 1-39 FROM N.A.
RX MEDLINE=91007285; PubMed=2170235;
RA Kawakami K., Yagawa Y., Nagano K.;
RT "Regulation of Na+,K+-ATPases. I. Cloning and analysis of the 5'-
RT flanking region of the rat NKAA2 gene encoding the alpha 2 subunit.";
RL Gene 91:267-270(1990).
CC -!- FUNCTION: THIS IS THE CATALYTIC COMPONENT OF THE ACTIVE ENZYME
CC WHICH CATALYZES THE HYDROLYSIS OF ATP COUPLED WITH THE EXCHANGE OF
CC NA AND K IONS ACROSS THE PLASMA MEMBRANE. THIS ACTION CREATES THE
CC ELECTROCHEMICAL GRADIENT OF NA AND K, PROVIDING THE ENERGY FOR
CC ACTIVE TRANSPORT OF VARIOUS NUTRIENTS.
CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + Na(+) (In) + K(+) (Out) = ADP +
CC phosphate + Na(+) (Out) + K(+) (In).
CC -!- SUBUNIT: COMPOSED OF THREE SUBUNITS: ALPHA (CATALYTIC), BETA
CC AND GAMMA.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to the cation transport ATPases family (P-type
CC ATPases). Subfamily IIC.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to licenses@isb-sib.ch).

DR EMBL; M14512; AAA0776.1; -;
DR EMBL; D90043; BAA14102.1; -;
DR PIR; B24639; B24639.
DR PIR; I54059; I54059.
DR HSP; P04191; 1EUL.
DR InterPro; IPR001757; ATPase_E1-E2.
DR InterPro; IPR006069; Cation_ATPase.
DR InterPro; IPR006068; Cation_ATPase_C.
DR InterPro; IPR004014; Cation_ATPase_N.
DR InterPro; IPR005834; Hydrolase.
DR InterPro; IPR005775; Na/K_ATPase_alpha.
DR Pfam; PF00689; Cation_ATPase_C; 1.
DR Pfam; PF00690; Cation_ATPase_N; 1.
DR Pfam; PF00122; E1-E2_ATPase; 1.
DR Pfam; PF00702; Hydrolase; 1.
DR PRINTS; PR00119; CATATPASE.
DR PRINTS; PR00121; NAKATPASE.
DR TIGRams; TIGR01106; ATPase-IIC-X-K; 1.
DR TIGRams; TIGR01494; ATPase_P-type; 5.
DR PROSITE; PS00154; ATPase_E1_E2; 1.
KW Hydrolase; Sodium/potassium transport; Transmembrane; Phosphorylation;
KW Magnesium; Metal-binding; ATP-binding; Multigene family.
FT PROPEP 1 5
FT CHAIN 5 1020 SODIUM/POTASSIUM-TRANSPORTING ATPASE
FT FT 5 ALPHA-2 CHAIN.
FT DOMAIN 6 85 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 86 106 POTENTIAL.
FT DOMAIN 107 129 LUMENAL (POTENTIAL).
FT TRANSMEM 130 150 POTENTIAL.
FT DOMAIN 151 286 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 287 306 POTENTIAL.
FT DOMAIN 307 318 LUMENAL (POTENTIAL).
FT TRANSMEM 319 336 POTENTIAL.
FT DOMAIN 337 769 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 770 789 POTENTIAL.
FT DOMAIN 790 799 LUMENAL (POTENTIAL).
FT TRANSMEM 800 820 POTENTIAL.
FT DOMAIN 821 840 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 841 863 POTENTIAL.
FT DOMAIN 864 915 LUMENAL (POTENTIAL).
FT TRANSMEM 916 935 POTENTIAL.
FT DOMAIN 936 948 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 949 967 POTENTIAL.
FT DOMAIN 968 982 LUMENAL (POTENTIAL).
FT TRANSMEM 983 1003 POTENTIAL.
FT DOMAIN 1004 1020 CYTOPLASMIC (POTENTIAL).
FT MOD_RES 374 374
FT MOD_RES 940 940
FT BINDING 80 82
FT METAL 714 714
FT METAL 718 718
SQ SEQUENCE 1020 AA; 112217 MW; 5436E795BD5B4CFA CRC64;
Query Match 74.5%; Score 35; DB 1; Length 1020;
Best Local Similarity 56.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 STRAEGGGGR 9
DB 14 TTAENGCGK 22

RESULT 14
MYSS CYPCA STANDARD; PRT; 1935 AA.
AC Q90339;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Myosin heavy chain, fast skeletal muscle.
DE Cyprinus carpio (Common carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;
OC Cyprinidae; Cyprinus.
OX NCBI_TaxID=7962;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fast muscle;
RX MEDLINE=97352533; PubMed=9208928;
RA Hirayama Y., Watabe S.;
RT "Structural differences in the crossbridge head of temperature-
RT associated myosin subfragment-1 isoforms from carp fast skeletal
RT muscle.";
RL Eur. J. Biochem. 246:380-387(1997).
RN [2]
RP SEQUENCE OF 981-1935 FROM N.A.
RC TISSUE=Fast muscle;
RX MEDLINE=97176447; PubMed=9023993;
RA Imai J., Hirayama Y., Kikuchi K., Kakinuma M., Watabe S.;
RT "cDNA cloning of myosin heavy chain isoforms from carp fast skeletal
RT muscle and their gene expression associated with temperature
RT acclimation.";
RL J. Exp. Biol. 200:27-34(1997).
RN [3]
RP SEQUENCE OF 1387-1528 FROM N.A.
EX MEDLINE=95194396; PubMed=7887920;
RA Watabe S., Imai J., Nakaya M., Hirayama Y., Okamoto Y., Masaki H.,
RA Uozumi T., Hirohno I., Aoki T.;

SEQUENCE OF 1145-1270.
 Watanabe B.; PubMed=2713098;
 "Amino-acid sequence of the hinge region in chicken myosin
 subfragment-2.";
 Biol. Chem. Hoppe-Seyler 370:55-61 (1989).
 [10]
 SEQUENCE OF 1857-1938 FROM N.A.
 MEDLINE=87217964; PubMed=3034534;
 Moriarty D.M., Barringer K.J., Dodgson J.B., Richter H.E.,
 Young R.B.;
 "Genomic clones encoding chicken myosin heavy-chain genes.";
 DNA 6:91-99 (1987).
 [11]
 X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 1-843.
 MEDLINE=93303624; PubMed=8316857;
 Rayment I., Rypniewski W.R., Schmidt-Base K., Smith R.,
 Tomchick D.R., Benning M.M., Winkelmann D.A., Wesenberg G.,
 Holden H.M.;
 "Three-dimensional structure of myosin subfragment-1: a molecular
 motor.";
 Science 261:50-58 (1993).
 -!- FUNCTION: MUSCLE CONTRACTION. MYOSIN IS A PROTEIN THAT BINDS TO
 F-ACTIN AND HAS ATPASE ACTIVITY THAT IS ACTIVATED BY F-ACTIN.
 -!- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
 AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
 -!- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
 -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE. SHOWING
 CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
 CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
 -!- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
 ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
 -!- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
 MEROMYOSIN (LMW) AND 1 HEAVY MEROMYOSIN (HWM). IT CAN LATER BE
 SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
 SUBFRAGMENT (S2).
 -!- SIMILARITY: Contains 1 myosin-like globular head domain.
 -!- SIMILARITY: Contains 1 IQ domain.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL; U87231; AAB47555.1; -;
 EMBL; M16557; AAB48970.1; -;
 PDB; 2MYS; 11-JAN-97.
 PDB; 1ALM; 17-DEC-97.
 PDB; 1M8Q; 10-SEP-02.
 InterPro; IPR000048; IQ region.
 InterPro; IPR001609; myosin head.
 InterPro; IPR004009; myosin N.
 InterPro; IPR002928; Myosin_tail.
 Pfam; PF00612; IQ; 2.
 Pfam; PF00683; myosin_head; 1.
 Pfam; PF02736; Myosin_N; 1.
 Pfam; PF01576; Myosin_tail; 1.
 PRINTS; PR00193; MYOSINHEAVY.
 ProDom; PD000355; myosin_head; 1.
 SMART; SM00015; IQ; 1.
 SMART; SM00242; MYSC; 1.
 PROSITE; PS50096; IQ; 1.
 Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
 ATP-binding; Methylation; Alkylation; Phosphorylation; Acetylation;
 Calmodulin-binding; Multigene family; 3D-structure.
 INIT MET 0 0
 FT DOMAIN 1 782 MYOSIN HEAD-LIKE.
 FT DOMAIN 783 812 IQ.
 FT DOMAIN 838 840 HINGE.

FT	DOMAIN	841	1938	
FT	NP_BIND	179	186	COILED COIL (POTENTIAL).
FT	DOMAIN	657	679	ATP (POTENTIAL).
FT	DOMAIN	759	773	ACTIN-BINDING.
FT	MOD_RES	1	1	ACTIN-BINDING.
FT	MOD_RES	35	35	ACETYLATION.
FT	MOD_RES	130	130	METHYLATION (MONO-).
FT	MOD_RES	551	551	METHYLATION (TRI-).
FT	MOD_RES	755	755	METHYLATION (TRI-).
FT	MOD_RES	697	697	METHYLATION (MONO-).
FT	MOD_RES	707	707	ALKYLATION (SH-1).
FT	CONFLICT	907	907	ALKYLATION (SH-2).
FT	CONFLICT	980	980	C -> Q (IN REF. 7 AND 8).
FT	CONFLICT	1343	1343	L -> F (IN REF. 1).
FT	CONFLICT	1545	1545	E -> D (IN REF. 5).
FT	CONFLICT	1796	1796	S -> A (IN REF. 5).
FT	CONFLICT	1830	1830	HV -> QL (IN REF. 5).
FT	CONFLICT	1863	1863	S -> A (IN REF. 5).
FT	CONFLICT	1929	1929	I -> V (IN REF. 10).
FT	CONFLICT	1931	1931	IHG -> FH (IN REF. 10).
FT	TURN	7	8	
FT	TURN	11	12	
FT	HELEX	14	17	
FT	TURN	21	24	
FT	HELEX	25	28	
FT	TURN	34	36	
FT	STRAND	37	41	
FT	STRAND	47	55	
FT	STRAND	59	62	
FT	STRAND	70	73	
FT	HELEX	74	76	
FT	STRAND	77	78	
FT	TURN	83	84	
FT	STRAND	90	90	
FT	HELEX	91	93	
FT	HELEX	99	108	
FT	TURN	109	113	
FT	STRAND	116	118	
FT	STRAND	123	126	
FT	HELEX	133	135	
FT	TURN	137	138	
FT	HELEX	139	142	
FT	TURN	143	145	
FT	TURN	148	149	
FT	HELEX	155	169	
FT	TURN	170	170	
FT	STRAND	173	179	
FT	TURN	181	182	
FT	HELEX	185	199	
FT	TURN	200	200	
FT	HELEX	218	233	
FT	STRAND	234	235	
FT	STRAND	243	244	
FT	STRAND	247	254	

Query Match 74.5%; Score 35; DB 1; Length 1938;
 Best Local Similarity 85.7%; Pred. NO. 3.3e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 AEGGGGR 9
 |||||
 DB 630 AEGGGGR 636

Search completed: December 11, 2003, 18:21:11
 Job time : 12.6667 secs

RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Reil S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.,
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL939112; CAB63312.1; -.
DR InterPro; IPR003009; FMN enzyme.
DR InterPro; IPR004652; NifR3 YhdG.
DR InterPro; IPR001269; UPF0034.
DR Pfam; PF01207; Duf; 1; nifR3_yhdG; 1.
DR TIGRFAMs; TIGR00737; nifR3_yhdG; 1.
DR PROSITE; PS01136; UPF0034; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 406 AA; 43621 MW; 820CAFFQ2DC2C2C CRC64;
Query Match 85.1%; Score 40; DB 16; Length 406;
Best Local Similarity 77.8%; Pred. No. 67;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 STAEVGGGR 9
Db 300 ATVEGGGR 308
RESULT 2
Q9PWF4 PRELIMINARY; PRT; 549 AA.
AC Q9PWF4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE Transcription factor RelB.
GN RELB.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]_TaxID=9031;
RP SEQUENCE FROM N.A.
RC TISSUE=Splice;
RA Pfaff K.A., Ikeda T., Hrdlickova R., Nehya J., Liss A., Huang S.,
RA Sif S., Gilmore T.D., Bose H.R.;
RT "Characterization of the chicken RelB transcription factor.";
RL Submitted (Oct-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF029260; AAD41539.1; -.
DR HSSP; P25799; 1BFT.
DR InterPro; IPR002909; IPT_TIG.
DR InterPro; IPR000451; NF_Rel_dor_fam.
DR Pfam; PF00554; RHD; 1.
DR Pfam; PF01833; TIG; 1.
DR PRINTS; PRO0057; NFEBTNSCPFCT.
DR SMART; SM00429; IPT; 1.
DR PROSITE; PS01204; REL_1; 1.
DR PROSITE; PS00354; REL_2; 1.
SQ SEQUENCE 549 AA; 60207 MW; DAA762F34073DD7E CRC64;
Query Match 85.1%; Score 40; DB 13; Length 549;
Best Local Similarity 87.5%; Pred. No. 92;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 TAEVGGGR 9
Db 520 TSEVGGGR 527
RESULT 3
Q8C6J8 PRELIMINARY; PRT; 137 AA.
ID Q8C6J8

AC Q8C6J8;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_TaxID=10090;
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL 60,770 full-length cDNAs.";
DR EMBL; AK054434; BAC35777.1; -.
KW Hypothetical protein.
SQ SEQUENCE 137 AA; 14339 MW; 0CAB983720D6B1B8 CRC64;
Query Match 83.0%; Score 39; DB 11; Length 137;
Best Local Similarity 77.8%; Pred. No. 32;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 STAEVGGGR 9
Db 76 ATASGGGR 84
RESULT 4
Q9FRA0 PRELIMINARY; PRT; 198 AA.
AC Q9FRA0;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical 21.1 kDa protein.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]_TaxID=4530;
RP SEQUENCE FROM N.A.
RC Chow T.-Y., Hsing Y.-I.C., Chen H.-H., Wu H.-P., Chao Y.-T.,
RA Liu S.-M., Hsiao Y.-Y., Huang J.-J., Lee P.-F., Su C.-L., Chen C.-S.,
RA Shaw J.-P.;
RT "Oryza sativa PAC P0001A07 genomics sequence.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC084218; AAG48847.1; -.
DR Gramene; Q9FRA0; -.
KW Hypothetical protein.
SQ SEQUENCE 198 AA; 21077 MW; A394680F2D1D2486 CRC64;
Query Match 83.0%; Score 39; DB 10; Length 198;
Best Local Similarity 77.8%; Pred. No. 47;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 STAEVGGGR 9
Db 44 SSGEVGGGR 52
RESULT 5
Q9XEU0 PRELIMINARY; PRT; 247 AA.
ID Q9XEU0
AC Q9XEU0;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Zinc-finger protein 1.

```
RESULT 9
Q9FP02 PRELIMINARY; PRT; 298 AA.
AC Q9FP02;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE P0038C05.22 protein.
GN P0038C05.22.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, PAC
RT clone:P0038C05." to the EMBL/GenBank/DBJ databases.
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003044; BAB19349.1; -.
DR Gramene; Q9FP02; -.
SQ SEQUENCE 298 AA; 23663 MW; 7293976C04D0EC9C CRC64;

Query Match 80.9%; Score 38; DB 10; Length 298;
Best Local Similarity 77.8%; Pred. No. 1.1e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 STAEAGGGR 9
Db 184 ATAEGGGGR 192

RESULT 10
Q8LNU4 PRELIMINARY; PRT; 362 AA.
AC Q8LNU4;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN OSUNBA004IP03.11.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=33947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Kim M.M.,
RA Overton II L.L., Bera J.J., Tsitrin T., Krol M.I., Jarrahi B.B.,
RA Jin S.S., Koo H., Zismann V., Heiao J., Blunt S., Vanaken S.S.,
RA Uterback T.T., Zidblyum T.V., Wang Q.Q., Haas B.J., Suh B.B.,
RA Peterson J.J., Quackenbush J., White O., Salzberg S.L., Fraser C.M.;
RT "Oryza sativa chromosome 10 EAC OSUNBA004IP03 genomic sequence."
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC068950; AA952818.1; -.
DR Gramene; Q8LNU4; -.
KW Hypothetical protein.
SQ SEQUENCE 362 AA; 36904 MW; 7D8FBE90C1623DE7 CRC64;

Query Match 80.9%; Score 38; DB 10; Length 362;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 AEGGGGR 9
Db 153 AEGGGGR 159

RESULT 11
Q9ARP7 PRELIMINARY; PRT; 383 AA.
AC Q9ARP7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE OSUNBA0010K01.19 protein.
GN OSUNBA0010K01.19.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
RT clone:OSUNBA0010K01." to the EMBL/GenBank/DBJ databases.
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003210; BAB40104.1; -.
DR Gramene; Q9ARP7; -.
SQ SEQUENCE 383 AA; 40697 MW; D0336CA0423584CF CRC64;

Query Match 80.9%; Score 38; DB 10; Length 383;
Best Local Similarity 77.8%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 STAEAGGGR 9
Db 55 STVSGGGGR 63

RESULT 12
Q9U0W5 PRELIMINARY; PRT; 389 AA.
AC Q9U0W5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Hypothetical 42.7 kDa protein.
GN L7276.04.
OS Leishmania major.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Friedlin;
RA Zimmermann W., Wambutt R., Ivens A.C., Murphy L., Quail M.,
RA Rajadream M.A., Barrell B.G.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Friedlin;
RX MEDLINE=98146435; PubMed=9477341;
RX Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
RX Smith D.F.;
RT "A physical map of the Leishmania major Friedlin genome."
RL Genome Res. 8:135-145(1998).
DR EMBL; AL153436; CAB62821.1; -.
KW Hypothetical protein.
SQ SEQUENCE 389 AA; 42733 MW; 4C1A0CF31D0DC670 CRC64;

Query Match 80.9%; Score 38; DB 5; Length 389;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 AEGGGGR 9
Db 60 AEGGGGR 66

RESULT 13
```

DR SMART; SM00387; HATPase_c; 1.
 KW Complete proteome.
 SQ SEQUENCE 943 AA; 102259 MW; 3DCEDDED81FD30B6 CRC64;
 Query Match 80.9%; Score 38; DB 16; Length 943;
 Best Local Similarity 100.0%; Pred. No. 3.5e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 TAEGGGG 8
 Db 842 TAEGGGG 848

Search completed: December 11, 2003, 18:28:29
 Job time : 23.3333 secs

XX 18-NOV-1999.
 XX 12-MAY-1999; 99WO-US10549.
 XX 12-MAY-1998; 98US-0085197.
 PR 05-MAY-1999; 99US-0085197.
 XX (BOCK/) BOCK S C.
 PA (PICA/) PICARD V.
 PA (ZEND/) ZENDEHROUH P.
 XX Bock SC, Picard V, Zendehtrouh P;
 PI WPI; 2000-116274/10.
 XX New modified human antithrombin III compounds, used for treating e.g.
 PT sepsis, trauma, acute respiratory distress syndrome, restenosis,
 PT thrombosis, thromboembolism or stroke -
 XX Claim 13; Page 57; 75pp; English.
 XX The present sequence is from an antithrombin III (ATIII) variant, F2A'
 CC derived from human ATIII.N135A cDNA insert of the pBlueBac baculovirus
 CC expression construct and comprises residues 385-393. The variant has
 CC improved resistance to elastase and IGG-activated neutrophils while
 CC retains anti-thrombin and anti-factor Xa activities. It may be
 CC expressed as glycoforms with enhanced heparin affinity which target the
 CC blood vessel wall more efficiently than ATIIIs with normal heparin
 CC affinity. The modified ATIIIs can be used to treat thrombin activation-
 CC related pathological symptoms due to sepsis, trauma, acute
 CC respiratory distress syndrome, restenosis, thrombosis, thromboembolism
 CC and stroke. It can also be used to reduce the risk of reocclusion
 CC and restenosis in percutaneous transluminal coronary angioplasty,
 CC thrombosis associated with surgery, ischaemia/reperfusion injury, and
 CC coagulation abnormalities in cancer or surgical patients.
 XX Sequence 9 AA;
 SQ Query Match 100.0%; Score 50; DB 21; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 STQTPPNR 9
 DB 1 STQTPPNR 9
 RESULT 2
 ABG01822
 ID ABG01822 standard; Protein; 224 AA.
 XX AC ABG01822;
 XX 13-FEB-2002 (first entry)
 XX Novel human diagnostic protein #1813.
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX Homo sapiens.
 OS WO200175067-A2.
 PN 11-OCT-2001.
 XX 30-MAR-2001; 2001WO-US08631.
 XX 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX (HYSE-) HYSEQ INC.
 PA

XX Drmanac RT, Liu C, Tang YT;
 PI WPI; 2001-639362/73.
 XX N-PSDB; AAS66009.
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX Claim 20; SEQ ID No 32181; 103pp; English.
 XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (II) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG0377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 224 AA;
 SQ Query Match 74.0%; Score 37; DB 22; Length 224;
 Best Local Similarity 77.8%; Pred. No. 1.5e+02;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 STQTPPNR 9
 DB 172 STQSPNRR 180
 RESULT 3
 AAY04833
 ID AAY04833 standard; Protein; 99 AA.
 XX AC AAY04833;
 XX 06-JUL-1999 (first entry)
 XX Mycobacterium species protein sequence 15C.
 DE Secreted protein; Mycobacterium; primer; PCR; amplification; probe;
 KW hybridisation; detection; vaccine; immunisation; infection.
 XX Mycobacterium sp.
 OS WO9909186-A2.
 PN 25-FEB-1999.
 XX 14-AUG-1998; 98WO-FR01813.
 XX 11-SEP-1997; 97FR-0011325.
 PR 14-AUG-1997; 97FR-0010404.
 XX (INSP) INST PASTEUR.
 XX Gicquel B, Lim EM, Pelicic V, Portnoi D, Goguet de la Salmoniere Y;
 PI

```

XX WPI; 2001-061974/07.
DR N-PSDB; AAF71246.
XX
XX New isolated Corynebacterium glutamicum nucleic acid for production or
PT modulation of production of fine chemicals such as amino acids,
PT nucleosides, nucleotides, lipids, fatty acids, carbohydrates, vitamins
PT or enzymes -
XX
XX Claim 20; Page 434-435; 712pp; English.
XX
XX AAF71138 to AAF71137 encode the Corynebacterium glutamicum homeostasis
CC and adaptation (HA) proteins given in AAB79023 to AAB79242. The
CC C. glutamicum HA genes (1) can be used in vectors for expression in host
CC cells and production of fine chemicals, such as, an organic acid, or
CC proteinogenic or nonproteinogenic amino acid (preferred), purine or
CC pyrimidine base, nucleoside, nucleotide, lipid, saturated or unsaturated
CC fatty acid, diol, carbohydrate, aromatic compound, vitamin, cofactor,
CC polyketide or enzyme. The amino acids produced can be lysine, glutamine,
CC glutamate, alanine, leucine, aspartate, glycine, serine, threonine, methionine,
CC cysteine, valine, leucine, isoleucine, arginine, proline, histidine,
CC tyrosine, phenylalanine, or tryptophan. The fine chemical production can
CC be modulated. The presence of (1) or HA proteins encoded by then are
CC used for diagnosing the presence or activity of Corynebacterium
CC diphtheriae. (1) can be used to map the C. glutamicum genome or can be
CC used as markers for genetically engineered Corynebacterium or
CC Brevibacterium. The HA proteins encoded by the (1) are used to maintain
CC homeostasis in C. glutamicum or help the microorganism to adapt to
CC different environmental conditions.
XX
XX Sequence 322 AA;
SQ
Query Match 72.0%; Score 36; DB 22; Length 322;
Best Local Similarity 85.7%; Pred. No. 3.2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 3 QTFPNCGR 9
Db 314 QTFPDGR 320
|||||
RESULT 6
AAB79666
ID AAB79666 standard; Protein; 322 AA.
XX
XX AAB79666;
AC
XX 30-APR-2001 (first entry)
DT
XX Corynebacterium glutamicum MP protein sequence SEQ ID NO:66.
DE
XX
XX Corynebacterium glutamicum; metabolic pathway protein; MP protein;
KW fine chemical production; microorganism; organic acid; nucleoside;
KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleotide;
KW lipid; saturated fatty acid; unsaturated fatty acid; diol; vitamin;
KW carbohydrate; aromatic compound; cofactor; polyketide; enzyme.
XX
XX Corynebacterium glutamicum.
OS
XX
XX WO200100843-A2.
PN
XX 04-JAN-2001.
PD
XX
XX 23-JUN-2000; 2000WO-IB00923.
PF
XX
XX 25-JUN-1999; 99US-0141031.
PR
XX 01-JUL-1999; 99DE-1030476.
PR
XX 02-JUL-1999; 99US-0142101.
PR
XX 08-JUL-1999; 99DE-1031415.
PR
XX 08-JUL-1999; 99DE-1031418.
PR
XX 08-JUL-1999; 99DE-1031419.
PR
XX 08-JUL-1999; 99DE-1031420.
PR
XX 08-JUL-1999; 99DE-1031424.
PR

```

```

PR 08-JUL-1999; 99DE-1031428.
PR 08-JUL-1999; 99DE-1031434.
PR 08-JUL-1999; 99DE-1031435.
PR 08-JUL-1999; 99DE-1031443.
PR 08-JUL-1999; 99DE-1031453.
PR 08-JUL-1999; 99DE-1031457.
PR 08-JUL-1999; 99DE-1031465.
PR 08-JUL-1999; 99DE-1031478.
PR 08-JUL-1999; 99DE-1031510.
PR 08-JUL-1999; 99DE-1031541.
PR 08-JUL-1999; 99DE-1031573.
PR 08-JUL-1999; 99DE-1031592.
PR 08-JUL-1999; 99DE-1031632.
PR 08-JUL-1999; 99DE-1031634.
PR 08-JUL-1999; 99DE-1031636.
PR 09-JUL-1999; 99DE-1032125.
PR 09-JUL-1999; 99DE-1032126.
PR 09-JUL-1999; 99DE-1032130.
PR 09-JUL-1999; 99DE-1032186.
PR 09-JUL-1999; 99DE-1032206.
PR 09-JUL-1999; 99DE-1032227.
PR 09-JUL-1999; 99DE-1032228.
PR 09-JUL-1999; 99DE-1032229.
PR 09-JUL-1999; 99DE-1032230.
PR 14-JUL-1999; 99DE-1032322.
PR 14-JUL-1999; 99DE-1032326.
PR 14-JUL-1999; 99DE-1032928.
PR 14-JUL-1999; 99DE-1033004.
PR 14-JUL-1999; 99DE-1033005.
PR 14-JUL-1999; 99DE-1033006.
PR 12-AUG-1999; 99US-0148613.
PR 27-AUG-1999; 99DE-1040764.
PR 27-AUG-1999; 99DE-1040765.
PR 27-AUG-1999; 99DE-1040766.
PR 27-AUG-1999; 99DE-1040832.
PR 31-AUG-1999; 99DE-1041378.
PR 31-AUG-1999; 99DE-1041379.
PR 31-AUG-1999; 99DE-1041380.
PR 31-AUG-1999; 99DE-1041394.
PR 31-AUG-1999; 99DE-1041396.
PR 03-SEP-1999; 99DE-1042076.
PR 03-SEP-1999; 99DE-1042077.
PR 03-SEP-1999; 99DE-1042079.
PR 03-SEP-1999; 99DE-1042086.
PR 03-SEP-1999; 99DE-1042087.
PR 03-SEP-1999; 99DE-1042088.
PR 03-SEP-1999; 99DE-1042095.
PR 03-SEP-1999; 99DE-1042124.
PR 03-SEP-1999; 99DE-1042129.
PR 09-MAR-2000; 2000US-0187970.
XX
XX (BADI ) BASF AG.
XX
XX Pompejus M, Kroegeer B, Schroeder H, Zelder O, Haberhauer G;
PI WPI; 2001-137957/14.
XX N-PSDB; AAF71785.
XX
XX Nucleic acids from Corynebacterium glutamicum encoding metabolic
PT pathway proteins, useful for producing fine chemicals in
PT microorganisms, including organic acids, nonproteinogenic amino acids,
PT and purine and pyrimidine bases -
XX
XX Claim 20; Page 249-250; 1737pp; English.
XX
XX AAF71753 to AAF72330 encode the Corynebacterium glutamicum metabolic
CC pathway (MP) proteins given in AAB79634 to AAB80211. The C. glutamicum
CC MP nucleic acids are useful for the production of fine chemicals
CC in microorganisms, including organic acids, nonproteinogenic amino
CC acids, purine and pyrimidine bases, nucleosides, nucleotides, lipids,
CC saturated and unsaturated fatty acids, diols, carbohydrates, aromatic
CC compounds, vitamins, cofactors, polyketides and enzymes.
XX

```

PN	BP1033405-A2.	PR	01-JUL-1999;	99US-0142154.
XX		PR	02-JUL-1999;	99US-0142055.
PD	06-SEP-2000.	PR	06-JUL-1999;	99US-0142390.
XX		PR	08-JUL-1999;	99US-0142803.
PF	25-FEB-2000; 2000BP-0301439.	PR	09-JUL-1999;	99US-0142920.
XX		PR	12-JUL-1999;	99US-0142977.
XX		PR	13-JUL-1999;	99US-0143542.
PR	25-FEB-1999;	PR	14-JUL-1999;	99US-0143624.
PR	05-MAR-1999;	PR	15-JUL-1999;	99US-0144005.
PR	09-MAR-1999;	PR	16-JUL-1999;	99US-0144086.
PR	23-MAR-1999;	PR	19-JUL-1999;	99US-0144325.
PR	29-MAR-1999;	PR	19-JUL-1999;	99US-0144331.
PR	25-MAR-1999;	PR	19-JUL-1999;	99US-0144332.
PR	01-APR-1999;	PR	19-JUL-1999;	99US-0144333.
PR	06-APR-1999;	PR	19-JUL-1999;	99US-0144334.
PR	08-APR-1999;	PR	19-JUL-1999;	99US-0144335.
PR	16-APR-1999;	PR	19-JUL-1999;	99US-0144352.
PR	19-APR-1999;	PR	20-JUL-1999;	99US-0144632.
PR	21-APR-1999;	PR	20-JUL-1999;	99US-0144632.
PR	23-APR-1999;	PR	20-JUL-1999;	99US-0144884.
PR	23-APR-1999;	PR	21-JUL-1999;	99US-0144814.
PR	28-APR-1999;	PR	21-JUL-1999;	99US-0145086.
PR	30-APR-1999;	PR	21-JUL-1999;	99US-0145088.
PR	30-APR-1999;	PR	22-JUL-1999;	99US-0145085.
PR	30-APR-1999;	PR	22-JUL-1999;	99US-0145087.
PR	04-MAY-1999;	PR	22-JUL-1999;	99US-0145089.
PR	05-MAY-1999;	PR	22-JUL-1999;	99US-0145192.
PR	06-MAY-1999;	PR	23-JUL-1999;	99US-0145145.
PR	07-MAY-1999;	PR	23-JUL-1999;	99US-0145218.
PR	11-MAY-1999;	PR	23-JUL-1999;	99US-0145218.
PR	11-MAY-1999;	PR	23-JUL-1999;	99US-0145224.
PR	14-MAY-1999;	PR	26-JUL-1999;	99US-0145276.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145913.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145918.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145919.
PR	18-MAY-1999;	PR	28-JUL-1999;	99US-0145951.
PR	19-MAY-1999;	PR	02-AUG-1999;	99US-0146386.
PR	20-MAY-1999;	PR	02-AUG-1999;	99US-0146388.
PR	21-MAY-1999;	PR	02-AUG-1999;	99US-0146388.
PR	21-MAY-1999;	PR	02-AUG-1999;	99US-0146389.
PR	24-MAY-1999;	PR	03-AUG-1999;	99US-0147038.
PR	25-MAY-1999;	PR	04-AUG-1999;	99US-0147204.
PR	27-MAY-1999;	PR	04-AUG-1999;	99US-0147302.
PR	28-MAY-1999;	PR	05-AUG-1999;	99US-0147192.
PR	01-JUN-1999;	PR	05-AUG-1999;	99US-0147260.
PR	03-JUN-1999;	PR	06-AUG-1999;	99US-0147303.
PR	07-JUN-1999;	PR	06-AUG-1999;	99US-0147416.
PR	07-JUN-1999;	PR	09-AUG-1999;	99US-0147493.
PR	08-JUN-1999;	PR	09-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	PR	10-AUG-1999;	99US-0148171.
PR	10-JUN-1999;	PR	11-AUG-1999;	99US-0148319.
PR	14-JUN-1999;	PR	12-AUG-1999;	99US-0148341.
PR	16-JUN-1999;	PR	13-AUG-1999;	99US-0148565.
PR	16-JUN-1999;	PR	13-AUG-1999;	99US-0148684.
PR	17-JUN-1999;	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	PR	27-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	PR	27-AUG-1999;	99US-0151066.
PR	18-JUN-1999;	PR	27-AUG-1999;	99US-0151080.
PR	18-JUN-1999;	PR	30-AUG-1999;	99US-0151303.
PR	21-JUN-1999;	PR	31-AUG-1999;	99US-0151438.
PR	23-JUN-1999;	PR	01-SEP-1999;	99US-0151930.
PR	23-JUN-1999;	PR	07-SEP-1999;	99US-0152363.
PR	23-JUN-1999;	PR	10-SEP-1999;	99US-0153070.
PR	24-JUN-1999;	PR	13-SEP-1999;	99US-0153758.
PR	28-JUN-1999;	PR	15-SEP-1999;	99US-0154018.
PR	29-JUN-1999;	PR	16-SEP-1999;	99US-0154039.
PR	30-JUN-1999;	PR		
PR	01-JUL-1999;	PR		

```
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144633.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147208.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 18-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151067.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.

PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
```

Query Match 70.0%; Score 35; DB 21; Length 345;

Best Local Similarity 85.7%; Pred. No. 5e+02;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 STCTPPN 7

Db 2 STXTPPN 8

Search completed: December 11, 2003, 18:25:02

Job time : 27 secs

RESULT 2
 US-09-270-984A-2
 ; Sequence 2, Application US/09270984A
 ; Patent No. 6048965
 ; GENERAL INFORMATION:
 ; APPLICANT: Dietz, Harry C.
 ; TITLE OF INVENTION: MAMMALIAN REGULATOR OF
 ; TITLE OF INVENTION: NONSENSE-MEDIATED RNA DECAY
 ; NUMBER OF SEQUENCES: 29
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson, P.C.
 ; STREET: 4225 Executive Square, Suite 1400
 ; CITY: La Jolla
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 92037
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: Windows95
 ; SOFTWARE: FASTSEQ for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/270,984A
 ; FILING DATE:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/724,354
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Haile, Lisa A.
 ; REGISTRATION NUMBER: 38,347
 ; REFERENCE/DOCKET NUMBER: 07265/090001
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 619-678-5070
 ; TELEFAX: 619-678-5099
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1118 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-09-270-984A-2

Query Match 70.0%; Score 35; DB 3; Length 1118;
 Best Local Similarity 85.7%; Pred. No. 4.6e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TOTPPNG 8
 Db 44 TOTPPGG 50

RESULT 3
 US-08-875-533-65
 ; Sequence 65, Application US/08875533
 ; Patent No. 6254870
 ; GENERAL INFORMATION:
 ; APPLICANT:
 ; TITLE OF INVENTION: No. 6254870e1 c-MPL Ligands
 ; NUMBER OF SEQUENCES: 73
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/875,533
 ; FILING DATE:
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/383,035
 ; FILING DATE: 04-FEB-1995
 ; INFORMATION FOR SEQ ID NO: 65:

; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 153 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-875-533-65
 ; Query Match 58.0%; Score 34; DB 3; Length 153;
 ; Best Local Similarity 75.0%; Pred. No. 86;
 ; Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 ; Qy 2 TOTPPNGR 9
 ; Db 110 TOLPPQGR 117
 ; RESULT 4
 ; US-08-469-318-164
 ; Sequence 164, Application US/08469318
 ; Patent No. 6022535
 ; GENERAL INFORMATION:
 ; APPLICANT:
 ; TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis Fusion
 ; TITLE OF INVENTION: Protein
 ; NUMBER OF SEQUENCES: 196
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/469,318
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/446,872
 ; FILING DATE:
 ; INFORMATION FOR SEQ ID NO: 164:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 155 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-469-318-164

Query Match 68.0%; Score 34; DB 3; Length 155;
 Best Local Similarity 75.0%; Pred. No. 87;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TOTPPNGR 9
 Db 112 TOLPPQGR 119

RESULT 5
 US-08-468-609A-164
 ; Sequence 164, Application US/08468609A
 ; Patent No. 6030812
 ; GENERAL INFORMATION:
 ; APPLICANT: Adams, Mark A.
 ; APPLICANT: Bauer, S. C.
 ; APPLICANT: Braford-Goldberg, Sarah R.
 ; APPLICANT: Caparon, Mair H.
 ; APPLICANT: Easton, Alan M.
 ; APPLICANT: Klein, Barbara K.
 ; APPLICANT: McKearn, John P.
 ; APPLICANT: Orlins, Peter O.
 ; APPLICANT: Paik, Kuman
 ; APPLICANT: Thomas, John W.
 ; TITLE OF INVENTION: Fusion Proteins Comprising Multiply Mutated Interleukin-3 (IL-3)
 ; NUMBER OF SEQUENCES: 197

APPLICATION NUMBER: US 08/446,872
FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Bennett, Dennis A.
REGISTRATION NUMBER: 34,547
REFERENCE/DOCKET NUMBER: C-2790/5
TELECOMMUNICATION INFORMATION:
TELEPHONE: (708)470-6501
TELEFAX: (708)470-6881
INFORMATION FOR SEQ ID NO: 164:
SEQUENCE CHARACTERISTICS:
LENGTH: 155 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 164:
US-08-762-227A-164

Query Match 68.0%; Score 34; DB 4; Length 155;
Best Local Similarity 75.0%; Pred. No. 87;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TOTPPNGR 9
DB 112 TQLPPQGR 119

RESULT 8
PCT-US95-01185-164
Sequence 164, Application PC/TUS9501185
GENERAL INFORMATION:
APPLICANT: Multivariant IL-3 Hematopoiesis Fusion
TITLE OF INVENTION: Protein
NUMBER OF SEQUENCES: 196
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/01185
FILING DATE: 02-FEB-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/192325
FILING DATE: 14-FEB-1994
INFORMATION FOR SEQ ID NO: 164:
SEQUENCE CHARACTERISTICS:
LENGTH: 155 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-01185-164

Query Match 69.0%; Score 34; DB 5; Length 155;
Best Local Similarity 75.0%; Pred. No. 87;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TOTPPNGR 9
DB 112 TQLPPQGR 119

RESULT 9
US-08-413-803-29
Sequence 29, Application US/08413803
Patent No. 5765581
GENERAL INFORMATION:
APPLICANT: Bartley, Timothy D. M.
APPLICANT: Bogenberger, Jakob M.

APPLICANT: Bosselman, Robert A.
APPLICANT: Hunt, Pamela
APPLICANT: Kinstler, Olaf B.
APPLICANT: Samal, Babru B.
TITLE OF INVENTION: METHODS FOR TREATING MAMMALS WITH
REFERENCE/DOCKET NUMBER: MONO-REGULATED PROTEINS THAT STIMULATE MEGAKARYOCYTE
TITLE OF INVENTION: GROWTH AND DIFFERENTIATION
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: AMGEN INC.
STREET: 1840 DeHavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: US
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/413,803
FILING DATE: 30-MAR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/221,768
FILING DATE: 31-MAR-1994
APPLICATION NUMBER: US 08/252,628
FILING DATE: 31-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/321,488
FILING DATE: 12-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/347,780
FILING DATE: 30-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Cook Ph.D., Robert R.
REGISTRATION NUMBER: 31,602
REFERENCE/DOCKET NUMBER: A-290D
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 165 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-413-803-29

Query Match 68.0%; Score 34; DB 1; Length 165;
Best Local Similarity 75.0%; Pred. No. 93;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TOTPPNGR 9
DB 112 TQLPPQGR 119

RESULT 10
US-08-471-045-56
Sequence 56, Application US/08471045
Patent No. 6060047
GENERAL INFORMATION:
APPLICANT: Abrams, Mark A.
APPLICANT: Bauer, S. C.
APPLICANT: Braford-Goldberg, Sarah R.
APPLICANT: Caparon, Mairé H.
APPLICANT: Easton, Alan M. K.
APPLICANT: Klein, Barbara K.
APPLICANT: McKearn, John P.
APPLICANT: Oline, Peter O.
APPLICANT: Paik, Kuman
APPLICANT: Thomas, John W.
TITLE OF INVENTION: Co-administration of Interleukin-3

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,871
; FILING DATE: 06-JUN-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/01184
; FILING DATE: 02-FEB-1995
; PRIOR APPLICATION DATA: US 08/193,373
; FILING DATE: 04-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bennett, Dennis A.
; REGISTRATION NUMBER: 34,547
; REFERENCE/DOCKET NUMBER: C-2789/1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (708)470-6501
; TELEFAX: (708)470-6881
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 174 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-446-871-56

Query Match 68.0%; Score 34; DB 4; Length 174;
Best Local Similarity 75.0%; Pred. No. 98;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TOTPPNGR 9
DB 131 TQLPPQGR 138

RESULT 13
US-08-468-910-56
; Sequence 56, Application US/08468910
; Patent No. 6379662
; GENERAL INFORMATION:
; APPLICANT: Abrams, Mark A.
; APPLICANT: Bauer, S. C.
; APPLICANT: Braford-Goldberg, Sarah R.
; APPLICANT: Caparon, Mairé H.
; APPLICANT: Easton, Alan M.
; APPLICANT: Klein, Barbara K.
; APPLICANT: McKearn, John P.
; APPLICANT: Olin, Peter O.
; APPLICANT: Paik, Kumnan
; APPLICANT: Thomas, John W.
; TITLE OF INVENTION: Co-administration of Interleukin-3
; TITLE OF INVENTION: Mutant Polypeptides with CSP's for Multi-lineage
; TITLE OF INVENTION: Hematopoietic Cell Production
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis A. Bennett, G.D. Searle & Co.,
; ADDRESSEE: Corporate Patent Dept.
; STREET: P. O. Box 5110
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60680
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,910

```

```

; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA: PCT/US95/01184
; FILING DATE: 02-FEB-1995
; PRIOR APPLICATION DATA: US 08/193,373
; FILING DATE: 04-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bennett, Dennis A.
; REGISTRATION NUMBER: 34,547
; REFERENCE/DOCKET NUMBER: C-2789/4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (708)470-6501
; TELEFAX: (708)470-6881
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 174 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-468-910-56

Query Match 68.0%; Score 34; DB 4; Length 174;
Best Local Similarity 75.0%; Pred. No. 98;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TOTPPNGR 9
DB 131 TQLPPQGR 138

RESULT 14
US-08-761-907-56
; Sequence 56, Application US/08761907
; Patent No. 6413509
; GENERAL INFORMATION:
; APPLICANT: Abrams, Mark A.
; APPLICANT: Bauer, S. C.
; APPLICANT: Braford-Goldberg, Sarah R.
; APPLICANT: Caparon, Mairé H.
; APPLICANT: Easton, Alan M.
; APPLICANT: Klein, Barbara K.
; APPLICANT: McKearn, John P.
; APPLICANT: Olin, Peter O.
; APPLICANT: Paik, Kumnan
; APPLICANT: Thomas, John W.
; TITLE OF INVENTION: Methods of Ex-vivo Expansion of Hematopoietic Cells
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis A. Bennett, G.D. Searle & Co.,
; ADDRESSEE: Corporate Patent Dept.
; STREET: P. O. Box 5110
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60680
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/761,907
; FILING DATE: 09-DEC-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/193,373
; FILING DATE: 04-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/01184
; FILING DATE: 02-FEB-1995

```

Sequence 14899, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: HORIKAWA, JUN
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
PRIOR FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 14899
LENGTH: 478
TYPE: PRT
ORGANISM: Streptomyces avermitilis
US-10-156-761-14899

Query Match 74.0%; Score 37; DB 15; Length 478;
Best Local Similarity 77.8%; Pred. No. 2.8e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 STOTPPNGR 9
Db 201 STWTPQGR 209

RESULT 3
US-09-746-660A-64
Sequence 64, Application US/09746660A
Publication No. US20030049804A1
GENERAL INFORMATION:
APPLICANT: Pompeius, Markus
APPLICANT: Kroger, Burkhard
APPLICANT: Schroder, Hartwig
APPLICANT: Zelder, Oskar
APPLICANT: Haberhauer, Gregor
APPLICANT: Kim, Jun-Won
APPLICANT: Lee, Heung-Schick
APPLICANT: Hwang, Byung-Joon
TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING
FILE REFERENCE: BGI-121CP2
CURRENT APPLICATION NUMBER: US/09/746,660A
CURRENT FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 09/606740
PRIOR FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: 09/603124
PRIOR FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: 60/141031
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: 60/142101
PRIOR FILING DATE: 1999-07-02
PRIOR APPLICATION NUMBER: 60/148613
PRIOR FILING DATE: 1999-08-12
PRIOR APPLICATION NUMBER: 60/187970
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: DE 19931420.9
PRIOR FILING DATE: 1999-07-08
NUMBER OF SEQ ID NOS: 125
SOFTWARE: PatentIn Vers. 2.0
SEQ ID NO 64
LENGTH: 322
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-746-660A-64

Query Match 72.0%; Score 36; DB 11; Length 322;
Best Local Similarity 85.7%; Pred. No. 2.8e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 QTPPNGR 9
Db 314 QTPPDGR 320

RESULT 4
US-09-738-626-5661
Sequence 5661, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 5661
LENGTH: 718
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-5661

Query Match 72.0%; Score 36; DB 10; Length 718;
Best Local Similarity 85.7%; Pred. No. 6e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 QTPPNGR 9
Db 314 QTPPDGR 320

RESULT 5
US-09-938-901-8
Sequence 8, Application US/09938901
Publication No. US20030008291A1
GENERAL INFORMATION:
APPLICANT: Kuramitsu Seiki,
APPLICANT: Yokoyama Shigeyuki
TITLE OF INVENTION: GENE ENCODING DNA REPAIR ENZYME
FILE REFERENCE: PH-1261-US
CURRENT APPLICATION NUMBER: US/09/938,901
CURRENT FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: JF2001-47762
PRIOR FILING DATE: 2001-02-23
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 8
LENGTH: 978
TYPE: PRT
ORGANISM: Thermus thermophilus
US-09-938-901-8

Db 112 TOLPPQGR 119

RESULT 9

US-10-072-571-56

; Sequence 56, Application US/10072571

; Publication No. US20030194783A1

; GENERAL INFORMATION:

; APPLICANT: Abrams, Mark A.

; Bauer, S. C.

; Bradford-Goldberg, Sarah R.

; Caparon, Mairé H.

; Easton, Alan M.

; Klein, Barbara K.

; McKearn, John P.

; Olin, Peter O.

; Paik, Kumman

; Thomas, John W.

; TITLE OF INVENTION: Co-administration of Interleukin-3

; Mutant Polypeptides with CSF's for Multi-lineage

; Hematopoietic Cell Production

; NUMBER OF SEQUENCES: 58

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: S. Christopher Bauer, Pharmacia Corp.

; Corporate Patent Dept.

; STREET: 800 N. Lindbergh

; CITY: St. Louis

; STATE: Missouri

; COUNTRY: USA

; ZIP: 63167

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.25

; CURRENT APPLICATION DATA: US/10/072,571

; FILING DATE: 08-Feb-2002

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/446,871

; FILING DATE: 06-JUN-1995

; APPLICATION NUMBER: PCT/US95/01184

; FILING DATE: 02-FEB-1995

; APPLICATION NUMBER: US 08/193,373

; FILING DATE: 04-FEB-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Bauer, S. Christopher

; REGISTRATION NUMBER: 42,305

; REFERENCE/DOCKET NUMBER: C-2789/6

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (636)737-5452

; TELEFAX: (636)737-5452

; INFORMATION FOR SEQ ID NO: 56:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 174 amino acids

; TYPE: amino acid

; STRANDEDNESS: <Unknown>

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 56:

US-10-072-571-56

Query Match 68.0%; Score 34; DB 12; Length 174;
Best Local Similarity 75.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy

2 TOTPPNGR 9

|||||

131 TOLPPQGR 138

|||||

RESULT 10

US-10-083-446-165

; Sequence 165, Application US/10083446

; Publication No. US20030185790A1

; GENERAL INFORMATION:

; APPLICANT: Abrams, Mark A.

; Bauer, S. C.

; Bradford-Goldberg, Sarah R.

; Caparon, Mairé H.

; Easton, Alan M.

; Klein, Barbara K.

; McKearn, John P.

; Olin, Peter O.

; Paik, Kumman

; Thomas, John W.

; TITLE OF INVENTION: Methods Of Ex-Vivo Expansion Of Hematopoietic Cells

; Using Multivariant (IL-3) Hematopoiesis Chimera Proteins

; NUMBER OF SEQUENCES: 197

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: S. Christopher Bauer, Pharmacia Corporation

; Corporate Patent Dept., Mail Zone O4E

; STREET: 800 N. Lindbergh

; CITY: St. Louis

; STATE: Missouri

; COUNTRY: USA

; ZIP: 63167

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.25

; CURRENT APPLICATION DATA: US/10/083,446

; FILING DATE: 26-Feb-2002

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/762,227

; FILING DATE: 09-DEC-1996

; APPLICATION NUMBER: US 08/192,325

; FILING DATE: 14-FEB-1994

; APPLICATION NUMBER: US 08/446,872

; FILING DATE: 06-JUN-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: S. Christopher Bauer

; REGISTRATION NUMBER: 42,305

; REFERENCE/DOCKET NUMBER: C-2790/6

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (636)737-5452

; TELEFAX: (636)737-5452

; INFORMATION FOR SEQ ID NO: 165:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 286 amino acids

; TYPE: amino acid

; STRANDEDNESS: <Unknown>

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 165:

US-10-083-446-165

Query Match 58.0%; Score 34; DB 12; Length 286;
Best Local Similarity 75.0%; Pred. No. 5.2e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy

2 TOTPPNGR 9

|||||

243 TOLPPQGR 250

|||||

RESULT 11

US-10-083-446-166

; Sequence 166, Application US/10083446

; Publication No. US20030185790A1

; GENERAL INFORMATION:

; APPLICANT: Abrams, Mark A.

```

/ Using Multivariant (IL-3) Hematopoiesis Chimera Proteins
/
/ NUMBER OF SEQUENCES: 197
/ CORRESPONDENCE ADDRESS:
/ ADDRESS: S. Christopher Bauer, Pharmacia Corporation
/ Corporate Patent Dept., Mail Zone 04E
/ STREET: 800 N. Lindbergh
/ CITY: St. Louis
/ STATE: Missouri
/ COUNTRY: USA
/ ZIP: 63167
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA: US/10/083,446
/ FILING DATE: 26-Feb-2002
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/762,227
/ FILING DATE: 09-DEC-1996
/ APPLICATION NUMBER: US 08/192,325
/ FILING DATE: 14-FEB-1994
/ APPLICATION NUMBER: US 08/446,872
/ FILING DATE: 06-JUN-1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: S. Christopher Bauer
/ REGISTRATION NUMBER: 42,305
/ REFERENCE/DOCKET NUMBER: C-2790/6
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (636)737-6257
/ TELEFAX: (636)737-5452
/
/ INFORMATION FOR SEQ ID NO: 168:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 290 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: <Unknown>
/ TOPOLOGY: linear
/ MOLECULE TYPE: Protein
/ SEQUENCE DESCRIPTION: SEQ ID NO: 168:
US-10-083-446-168

Query Match 68.0%; Score 34; DB 12; Length 290;
Best Local Similarity 75.0%; Pred. No. 5.3e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TOTPPNGR 9
Db 112 TQLPPQGR 119

RESULT 14
US-10-400-377-7
; Sequence 7, Application US/10400377
; Publication No. US20030162949A1
; GENERAL INFORMATION:
; APPLICANT: Cox III, George N
; TITLE OF INVENTION: Derivatives of Growth Hormone and Related Proteins
; FILE REFERENCE: 4152-1-PUS
; CURRENT APPLICATION NUMBER: US/10/400,377
; CURRENT FILING DATE: 2003-03-26
; PRIOR APPLICATION NUMBER: US/09/462,941
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: 60/052,516
; PRIOR FILING DATE: 1997-07-14
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 7
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-083-446-168

Query Match 68.0%; Score 34; DB 12; Length 290;
Best Local Similarity 75.0%; Pred. No. 5.3e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TOTPPNGR 9
Db 112 TQLPPQGR 119

RESULT 14
US-10-400-377-7
; Sequence 7, Application US/10400377
; Publication No. US20030162949A1
; GENERAL INFORMATION:
; APPLICANT: Cox III, George N
; TITLE OF INVENTION: Derivatives of Growth Hormone and Related Proteins
; FILE REFERENCE: 4152-1-PUS
; CURRENT APPLICATION NUMBER: US/10/400,377
; CURRENT FILING DATE: 2003-03-26
; PRIOR APPLICATION NUMBER: US/09/462,941
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: 60/052,516
; PRIOR FILING DATE: 1997-07-14
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 7
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-083-446-168

```

```

US-10-400-377-7
Query Match 68.0%; Score 34; DB 12; Length 332;
Best Local Similarity 75.0%; Pred. No. 6e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TOTPPNGR 9
Db 110 TQLPPQGR 117

RESULT 15
US-10-400-708-7
; Sequence 7, Application US/10400708
; Publication No. US2003016685A1
; GENERAL INFORMATION:
; APPLICANT: Bolder Biotechnology, Inc.
; TITLE OF INVENTION: Derivatives of Growth Hormone and Related Proteins
; FILE REFERENCE: 4152-1-PUS
; CURRENT APPLICATION NUMBER: US/10/400,708
; CURRENT FILING DATE: 2003-03-26
; PRIOR APPLICATION NUMBER: US/09/462,941
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: 60/052,516
; PRIOR FILING DATE: 1997-07-14
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 7
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-400-708-7

Query Match 68.0%; Score 34; DB 12; Length 332;
Best Local Similarity 75.0%; Pred. No. 6e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TOTPPNGR 9
Db 110 TQLPPQGR 117

Search completed: December 11, 2003, 18:38:05
Job time : 50.3333 secs

```

Query Match 72.0%; Score 36; DB 2; Length 290;
Best Local Similarity 66.7%; Pred. No. 33;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 STQTPNGR 9
| : |||||
DB 208 SLETPPDGR 216

RESULT 3
F72501
translation releasing factor arf-1 APE1988 [similarity] - Aeropyrum pernix (strain K1)
N;Alternate names: peptide chain release factor 1
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 15-Sep-2000
C;Accession: F72501
R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Maikawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum pernix
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: F72501
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-341 <RAW>
A;Cross-references: DDBJ:AF000063; NID:95105654; PIDN:BAA80998.1; PID:G5105686
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE1988
C;Superfamily: cell division protein MJ0174

Query Match 72.0%; Score 36; DB 2; Length 341;
Best Local Similarity 85.7%; Pred. No. 40;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TQTPNG 8
| : |||||
DB 54 TSITPPNG 60

RESULT 4
A55962
opsin, pineal gland-specific - chicken
N;Alternate names: P-opsin; pincopsin
C;Species: Gallus gallus (Chicken)
C;Date: 23-Mar-1995 #sequence_revision 05-Apr-1995 #text_change 22-Jun-1999
C;Accession: A55962; S50857
R;Max, M.; McKinnon, P.J.; Seidenman, K.J.; Barrett, R.K.; Applebury, M.L.; Takahashi, J Science 267, 1502-1506, 1995
A;Title: Pineal opsin: a nonvisual opsin expressed in chick pineal.
A;Reference number: A55962; MUID:95184012; PMID:7878470
A;Accession: A55962
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-351 <MAX>
R;Okano, T.; Yoshizawa, T.; Fukada, Y.
Nature 372, 94-97, 1994
A;Title: Pinopsin is a chicken pineal photoreceptive molecule.
A;Reference number: S50857; MUID:95059405; PMID:7969427
A;Accession: S50857
A;Molecule type: mRNA
A;Residues: 1-134, 'R', 136-162, 'T', 164-351 <OKA>
A;Cross-references: GB:U15762; NID:G726462; PIDN:AAA64223.1; PID:G726463
C;Function:
A;Description: modulation of melatonin biosynthesis by suppressing N-acetyltransferase a
A;Note: abundant in pineal gland but absent from retina
C;Superfamily: vertebrate rhodopsin
C;Keywords: chromoprotein; G protein-coupled receptor; glycoprotein; lipoprotein; phosph
F;1-30/Domain: extracellular #status predicted <EX1>
F;31-55/Domain: transmembrane #status predicted <TM1>
F;56-67/Domain: intracellular #status predicted <IN1>

F;68-90/Domain: transmembrane #status predicted <EX2>
F;91-107/Domain: extracellular #status predicted <EX2>
F;108-127/Domain: transmembrane #status predicted <TM3>
F;128-146/Domain: intracellular #status predicted <IN4>
F;147-169/Domain: transmembrane #status predicted <TM4>
F;170-194/Domain: extracellular #status predicted <EX3>
F;195-222/Domain: transmembrane #status predicted <TM5>
F;223-244/Domain: intracellular #status predicted <IN3>
F;245-268/Domain: transmembrane #status predicted <TM6>
F;269-277/Domain: extracellular #status predicted <EX4>
F;278-301/Domain: transmembrane #status predicted <TM7>
F;302-351/Domain: intracellular #status predicted <IN4>
F;104-181/Disulfide bonds: #status predicted
F;288/Binding site: retinal (lys) (covalent) #status predicted
F;314,315/Binding site: palmitate (Cys) (covalent) #status predicted

Query Match 72.0%; Score 36; DB 1; Length 351;
Best Local Similarity 75.0%; Pred. No. 41;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 STQTPNG 8
| : |||||
DB 5 SSQAPPNG 12

RESULT 5
S13237
tail fiber protein gp37 - phage T4 (fragment)
C;Species: Phage T4
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 20-Sep-1999
C;Accession: S13237
R;Montag, D.; Hahemolhosseini, S.; Henning, U.
J. Mol. Biol. 216, 327-334, 1990
A;Title: Receptor-recognizing proteins of T-even type bacteriophages. The receptor-reco
A;Reference number: S13237; MUID:91073397; PMID:2147721
A;Accession: S13237
A;Molecule type: DNA
A;Residues: 1-382 <MON>
A;Cross-references: EMBL:X55190; NID:G14860; PIDN:CAA38973.1; PID:G14861
A;Note: the sequence from Fig. 5 is inconsistent with that from Fig. 2 in having 25-Phe
C;Genetics:
A;Gene: gp37
C;Superfamily: phage T4 tail fiber protein gp37
C;Keywords: tail fiber

Query Match 72.0%; Score 36; DB 2; Length 382;
Best Local Similarity 85.7%; Pred. No. 45;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TQTPNG 8
| : |||||
DB 142 TDTTPNG 148

RESULT 6
T07280
photosystem I P700 apoprotein 1B - Chlorella vulgaris chloroplast
C;Species: chloroplast Chlorella vulgaris
C;Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 21-Jul-2000
C;Accession: T07280
R;Wakasugi, T.; Nagai, T.; Kapoor, M.; Sugita, M.; Ito, M.; Ito, S.; Tsudzuki, J.; Naka
Proc. Natl. Acad. Sci. U.S.A. 94, 5967-5972, 1997
A;Title: Complete nucleotide sequence of the chloroplast genome from the green alga Chl
A;Reference number: Z15985; MUID:97303241; PMID:9159184
A;Accession: T07280
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-734 <WAK>
A;Cross-references: EMBL:AB001684; NID:G2224352; PIDN:BAA57928.1; PID:G2224444
C;Genetics:
A;Gene: psab
C;Superfamily: photosystem I P700 apoprotein

C;Genetics:
A;Gene: P53A3.6
A;Map position: 3

Query Match 70.0%; Score 35; DB 2; Length 451;
Best Local Similarity 62.5%; Pred. No. 81;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 STOTPPNG 8
: : : : :
Db 216 ASETPPNG 223

RESULT 12

H84640
probable Rieske iron-sulfur protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cross)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: H84640
R;Lin, X.; Kaul, S.; Rounseley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayan, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: H84640
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-539 <STO>
A;Cross-references: GB:AE002093; NID:g4559369; PIDN:AAD23030.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2G24820
A;Map position: 2

Query Match 70.0%; Score 35; DB 2; Length 539;
Best Local Similarity 85.7%; Pred. No. 98;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 STOTPPN 7
: : : : :
Db 196 STKTPPN 202

RESULT 13

A2LVPF7
Photosystem I P700 apoprotein A2 - liverwort (Marchantia polymorpha) chloroplast
C;Species: chloroplast Marchantia polymorpha
C;Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1993
C;Accession: A03467; S01605
R;Ohnaya, K.
submitted to the EMBL Data Library, October 1986
A;Reference number: A00150
A;Accession: A03467
A;Molecule type: DNA
A;Residues: 1-734 <OHY>
R;Ohnaya, K.; Fukuzawa, H.; Kohchi, T.; Shirai, T.; Sano, S.; Umesono, K.; Shi
Nature 322, 572-574, 1986
A;Title: Chloroplast gene organization deduced from complete sequence of liverwort March
A;Reference number: A38014
A;Contents: annotation; Gene organization, sites, features
R;Umesono, K.; Inokuchi, H.; Shiki, Y.; Takeuchi, M.; Chang, Z.; Fukuzawa, H.; Kohchi, T
J. Mol. Biol. 203, 299-331, 1988
A;Title: Structure and organization of Marchantia polymorpha chloroplast genome. II. Gen
A;Reference number: S01567; MUID:89068686; PMID:2974085
A;Accession: S01605
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-734 <UNE>
A;Cross-references: GB:X04465; GB:Y00686; NID:g11640; PIDN:CAA28084.1; PID:g11671
C;Comment: This is one of the specific proteins associated with chlorophyll a of the P70
supplied by Photosystem II to Z-substance (an iron-sulfur protein), which in turn redu
lifer proteins.
C;Genetics:

A;Gene: psab
A;Genome: chloroplast
C;Superfamily: photosystem I P700 apoprotein
C;Keywords: chloroplast; electron transfer; membrane protein; membrane-associated compl.

Query Match 70.0%; Score 35; DB 1; Length 734;
Best Local Similarity 75.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TOTPPNGR 9
: : : : :
Db 307 THTPPGR 314

RESULT 14

B87681
tyrosine kinase DivL [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C;Accession: B87681
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolo
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: B87681
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-769 <STO>
A;Cross-references: GB:AE005673; NID:g13425208; PIDN:AAK25446.1; GSPDB:GN00148
C;Genetics:
A;Gene: CC3484

Query Match 70.0%; Score 35; DB 2; Length 769;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 QTPPNGR 9
: : : : :
Db 671 QTPPGR 677

RESULT 15

T50956
hypothetical protein B24P7.110 [imported] - Neurospora crassa
C;Species: Neurospora crassa
C;Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
C;Accession: T50956
R;Schulte, U.; Aign, V.; Hohnsels, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura
submitted to the Protein Sequence Database, July 2000
A;Reference number: Z25286
A;Accession: T50956
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-979 <SCH>
A;Cross-references: EMBL:AL389890; GSPDB:GN00116; NCSP:B24P7.110
A;Experimental source: BAC clone B24P7; strain OR74A
C;Genetics:
A;Gene: NCSP:B24P7.110
A;Map position: 6

Query Match 70.0%; Score 35; DB 2; Length 979;
Best Local Similarity 85.7%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 QTPPNGR 9
: : : : :
Db 718 QTPPNSR 724

Search completed: December 11, 2003, 18:29:44
Job time : 10.3333 secs

C/Genetics:

A:Gene: ATSP:T2711.1
 A:Map position: 1
 A:Introns: 58/2, 87/1, 138/3, 223/2, 241/3, 321/3, 366/2, 427/3, 447/3, 540/3, 627/1, 74/1
 A:Superfamily: Arabidopsis thaliana hypothetical protein T2711.1

Query Match 72.7%; Score 32; DB 2; Length 837;
 Best Local Similarity 66.7%; Pred. NO. 1.7e+02;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 STEVEGAGR 9

DB 305 SVEVEGSGQ 313

RESULT 13

AE2448
 hypothetical protein alr5141 [imported] - Nostoc sp. (strain PCC 7120)

C/Species: Nostoc sp. PCC 7120
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C/Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002

C/Accession: AE2448

R;Kakazaki, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Kaneko, T.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AE2448

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-110 <KUR>

A:Cross-references: GB:BA000019; PIDN:BA076840.1; PID:g17134279; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C/Genetics:

A:Gene: alr5141

Query Match 70.5%; Score 31; DB 2; Length 110;

Best Local Similarity 87.5%; Pred. NO. 33;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 STEVEGAG 8

DB 88 STEVETAG 95

RESULT 14

VXECSE

preprotein translocase secE chain - Escherichia coli (strain K-12)

C/Species: Escherichia coli

C/Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 01-Mar-2002

C/Accession: A35139; A32873; H65204

R;Downing, W.L.; Sullivan, S.L.; Gottesman, M.E.; Dennis, P.P.

J. Bacteriol. 172, 1621-1627, 1990

A:Title: Sequence and transcriptional pattern of the essential Escherichia coli secE-nus

A:Reference number: A35139; MUID:90170882; PMID:2137819

A:Accession: A35139

A:Molecule type: DNA

A:Residues: 1-127 <DOW>

A:Cross-references: GB:M30610; NID:g147798; PIDN:AAA24621.1; PID:g147800

R;Schatz, P.J.; Riggs, P.D.; Jacq, A.; Fath, M.J.; Beckwith, J.

Genes Dev. 3, 1035-1044, 1989

A:Title: The secE gene encodes an integral membrane protein required for protein export

A:Reference number: A32873; MUID:89378734; PMID:2673920

A:Accession: A32873

A:Molecule type: DNA

A:Residues: 1-127 <SCH>

R;Plattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; CC

.A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617; PMID:9278503

A:Accession: H65204

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-127 <BLAT>
 A:Cross-references: GB:AE000472; GB:U00096; NID:g2367333; PIDN:AA076955.1; PID:g1790413
 A:Experimental source: strain K-12, substrain MG1655

C/Comment: This integral inner membrane protein is an essential component of the protein

C/Genetics:

A:Gene: secE

A:Map position: 90 min

C:Superfamily: protein-export protein secE

C/Keywords: inner membrane; protein export

Query Match 70.5%; Score 31; DB 1; Length 127;

Best Local Similarity 55.6%; Pred. No. 39;

Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 STEVEGAGR 9

DB 4 NTEAQSGGR 12

RESULT 15

H91241

preprotein translocase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 05

C/Species: Escherichia coli

C/Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001

C/Accession: H91241

R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.C

Sasawara, N.; Yasunaga, T.; Khara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and ger

A:Reference number: A99629; MUID:21156231; PMID:11258796

A:Accession: H91241

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-127 <HAY>

A:Cross-references: GB:BA000007; PIDN:BA038327.1; PID:g13364380; GSPDB:GN00154

A:Experimental source: strain O157:H7, substrain RIMD 0509952

C/Genetics:

A:Gene: ECs4904

C:Superfamily: protein-export protein secE

Query Match 70.5%; Score 31; DB 2; Length 127;

Best Local Similarity 55.6%; Pred. No. 39;

Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 STEVEGAGR 9

DB 4 NTEAQSGGR 12

Search completed: December 11, 2003, 18:29:39

Job time : 10.3333 secs

Query Match 77.3%; Score 34; DB 1; Length 197;
 Best Local Similarity 75.0%; Pred. No. 14;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TEVEGAGR 9
 |:|:|:|
 Db 69 TQVDGAGR 76

RESULT 3
 T35032
 probable hydroxylase - Streptomyces coelicolor
 C/Species: Streptomyces coelicolor
 C/Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 21-Jan-2000
 C/Accession: T35032
 R/Seeger, S.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
 Submitted to the EMBL Data Library, June 1999
 A/Reference number: Z21565
 A/Accession: T35032
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-263 <SEE>
 A/Cross-references: EMBL:AL079355; PIDN:CAB45588.1; GSPDB:GN00070; SCOEDB:SC4C6.24C
 A/Experimental source: strain A3(2)
 C/Genetics:
 A/Gene: SCOEDB:SC4C6.24C
 C/Superfamily: Mycobacterium hypothetical protein Rv0911

Query Match 75.0%; Score 33; DB 2; Length 263;
 Best Local Similarity 55.6%; Pred. No. 31;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 STEVEGAGR 9
 |:|:|:|
 Db 233 ATDIEGVR 241

RESULT 4
 T11751
 transcription repressor ROM1 - kidney bean
 N/Alternate names: G-box binding factor
 C/Species: Phaseolus vulgaris (kidney bean)
 C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 18-Aug-2000
 C/Accession: T11751
 R/Chern, M.S.; Eiben, H.G.; Bustos, M.M.
 Plant J. 10, 135-148, 1996
 A/Title: The developmentally regulated bZIP factor ROM1 modulates transcription from leucine zipper
 A/Reference number: Z17333; MUID:96314865; PMID:8758983
 A/Accession: T11751
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-339 <CHE>
 A/Cross-references: EMBL:U57389; NID:g1354856; PIDN:AAB36514.1; PID:g1354857
 C/Function:
 A/Description: modulates transcription from lectin and storage protein genes
 C/Superfamily: fava bean G-box-binding protein; fos/jun DNA-binding domain homology
 C/Keywords: DNA binding; leucine zipper; transcription regulation

Query Match 75.0%; Score 33; DB 2; Length 339;
 Best Local Similarity 75.0%; Pred. No. 41;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 STEVEGAGR 8
 |:|:|:|
 Db 104 STEIEGRK 111

RESULT 5
 AD2536
 DNA polymerase III, delta prime chain [imported] - Nostoc sp. (strain PCC 7120) plasmid
 C/Species: Nostoc sp. PCC 7120
 A/Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C/Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002

C/Accession: AD2536
 R/Xaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguch, N.; Shimpou, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, DNA Res. 8, 205-213, 2001
 A/Title: Complete genomic sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ar
 A/Reference number: AB1807; MUID:21595285; PMID:11759840
 A/Accession: AD2536
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-291 <KUR>
 A/Cross-references: GB:AP003602; PIDN:BAW77218.1; PID:g17134660; GSPDB:GN00181
 A/Experimental source: strain PCC 7120
 C/Genetics:
 A/Gene: alr7575
 A/Genome: plasmid

Query Match 72.7%; Score 32; DB 2; Length 291;
 Best Local Similarity 66.7%; Pred. No. 57;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 STEVEGAGR 9
 |:|:|:|
 Db 35 SNEVEGVGK 43

RESULT 6
 AC0828
 Probable oxidoreductase STV2819 [imported] - Salmonella enterica subsp. enterica serovar
 C/Species: Salmonella enterica subsp. enterica serovar typhi
 A/Note: this species has also been called Salmonella typhi
 C/Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C/Accession: AC0828
 R/Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrah, S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A/Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.
 A/Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar
 A/Reference number: AB0502; MUID:21534947; PMID:11677608
 A/Accession: AC0828
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-305 <PAR>
 A/Cross-references: GB:AL513382; PIDN:CAD02775.1; PID:g16503785; GSPDB:GN00176
 C/Genetics:
 A/Gene: STV2819

Query Match 72.7%; Score 32; DB 2; Length 305;
 Best Local Similarity 75.0%; Pred. No. 60;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 STEVEGAGR 8
 |:|:|:|
 Db 128 SSELEGAG 135

RESULT 7
 F71243
 probable translation initiation factor eIF-2B delta - Pyrococcus horikoshii
 C/Species: Pyrococcus horikoshii
 C/Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 21-Jul-2000
 C/Accession: F71243
 R/Kawarayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sek, M.; Ohnuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi, DNA Res. 5, 55-76, 1998
 A/Title: Complete sequence and gene organization of the genome of a hyper-thermophilic
 A/Reference number: A71000; MUID:98344137; PMID:9679194
 A/Accession: F71243
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-324 <KAW>
 A/Cross-references: GB:AP000001; NID:g3236128; PIDN:BAA29277.1; PID:g3256594
 A/Experimental source: strain OT3

```

; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 26 Human secreted proteins
; FILE REFERENCE: P2040P1
; CURRENT APPLICATION NUMBER: US/10/042,141
; CURRENT FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: 09/726,643
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: PCT/US00/15187
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/137,725
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 144
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-042-141-144

Query Match          70.5%; Score 31; DB 14; Length 42;
Best Local Similarity 66.7%; Pred. No. 54;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 STEVEGAGR 9
      |||: |||
Db      30 SSEVDAAGR 38

RESULT 11
US-09-864-761-33858
; Sequence 33858, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeonica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30

```

```

; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33858
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007707.10
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 3.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.1
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.8
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.5
; OTHER INFORMATION: SWISSPROT HIT: 075312, EVALUE 3.00e-26
; OTHER INFORMATION: EST_HUMAN HIT: BE385905.1, EVALUE 3.00e-25
US-09-864-761-33858

```

```

Query Match          70.5%; Score 31; DB 9; Length 54;
Best Local Similarity 55.6%; Pred. No. 70;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1 STEVEGAGR 9
      |||: |||
Db      30 NTEIQSAGR 38

RESULT 12
US-09-815-242-10432
; Sequence 10432, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Olsens, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10432
; LENGTH: 127

```

; TITLE OF INVENTION: OF THE MU-OPIOID RECEPTOR GENE
; FILE REFERENCE: 830002-2000.3
; CURRENT APPLICATION NUMBER: US/10/283,300
; CURRENT FILING DATE: 2002-10-29
; PRIOR APPLICATION NUMBER: 09/761,962
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/743,872
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: PCT/US99/15974
; PRIOR FILING DATE: 1999-07-15
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-283-300-18

Query Match 81.8%; Score 36; DB 15; Length 359;
Best Local Similarity 87.5%; Pred. No. 55;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 STEVEGAG 8
Db 19 STELEGAG 26

RESULT 3

US-09-761-962-21
; Sequence 21, Application US/09761962
; Patent No. US2002007285A1
; GENERAL INFORMATION:
; APPLICANT: Memorial Sloan-Kettering Cancer Center
; TITLE OF INVENTION: Identification and Characterization of Multiple Splice
; TITLE OF INVENTION: Variants of Mu-
; TITLE OF INVENTION: Opioid Receptor (MOR-1) Gene
; FILE REFERENCE: 830002-2000.1
; CURRENT APPLICATION NUMBER: US/09/761,962
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/743,872
; PRIOR FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 21
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-761-962-21

Query Match 81.8%; Score 36; DB 9; Length 399;
Best Local Similarity 87.5%; Pred. No. 61;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 STEVEGAG 8
Db 19 STELEGAG 26

RESULT 4

US-10-283-300-21
; Sequence 21, Application US/10283300
; Publication No. US20030103972A1
; GENERAL INFORMATION:
; APPLICANT: Memorial Sloan-Kettering Cancer Center
; TITLE OF INVENTION: IDENTIFICATION AND CHARACTERIZATION OF MULTIPLE SPLICE VARIANTS
; TITLE OF INVENTION: OF THE MU-OPIOID RECEPTOR GENE
; FILE REFERENCE: 830002-2000.3
; CURRENT APPLICATION NUMBER: US/10/283,300
; CURRENT FILING DATE: 2002-10-29
; PRIOR APPLICATION NUMBER: 09/761,962
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/743,872
; PRIOR FILING DATE: 2001-01-16

; PRIOR APPLICATION NUMBER: PCT/US99/15974
; PRIOR FILING DATE: 1999-07-15
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 21
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-283-300-21

Query Match 81.8%; Score 36; DB 15; Length 399;
Best Local Similarity 87.5%; Pred. No. 61;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 STEVEGAG 8
Db 19 STELEGAG 26

RESULT 5

US-09-738-626-4481
; Sequence 4481, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4481
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4481

Query Match 77.3%; Score 34; DB 10; Length 135;
Best Local Similarity 87.5%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 STEVEGAG 8
Db 48 STEVEGLG 55

RESULT 6

US-10-156-761-13895
; Sequence 13895, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHISA

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 542-5070
 TELEFAX: (617) 542-8906
 TELEX: 200154
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 405
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: N/A
 US-08-470-925-2

Query Match 70.5%; Score 31; DB 2; Length 405;
 Best Local Similarity 75.0%; Pred. No. 2e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 STEVEGAG 8
 Db 255 STRVEGSG 262

RESULT 15
 US-08-471-613-2
 : Sequence 2, Application US/08471613
 : Patent No. 5962256
 : GENERAL INFORMATION:
 : APPLICANT: David D. Moore
 : APPLICANT: Jae Woon Lee
 : TITLE OF INVENTION: NUCLEAR HORMONE RECEPTOR-
 : TITLE OF INVENTION: INTERACTING POLYPEPTIDES AND
 : TITLE OF INVENTION: RELATED MOLECULES AND METHODS
 : NUMBER OF SEQUENCES: 31
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Fish & Richardson P.C.
 : STREET: 225 Franklin Street
 : CITY: Boston
 : STATE: Massachusetts
 : COUNTRY: U.S.A.
 : ZIP: 02110-2804
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 : COMPUTER: IBM PS/2 Model 502 or 55SX
 : OPERATING SYSTEM: MS-DOS (Version 5.0)
 : SOFTWARE: WordPerfect (Version 5.1)
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/471,613
 : FILING DATE: 06-June-1995
 : CLASSIFICATION: 436
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 08/222,719
 : FILING DATE: 04-April-1994
 : CLASSIFICATION: 436
 : APPLICATION NUMBER: 07/969,136
 : FILING DATE: October 30, 1992
 : CLASSIFICATION: 436
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Paul T. Clark
 : REGISTRATION NUMBER: 30,162
 : REFERENCE/DOCKET NUMBER: 00786/229001
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (617) 542-5070
 : TELEFAX: (617) 542-8906
 : TELEX: 200154
 : INFORMATION FOR SEQ ID NO: 2:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 405
 : TYPE: amino acid
 : STRANDEDNESS:
 : TOPOLOGY: N/A
 : US-08-471-613-2

Query Match 70.5%; Score 31; DB 2; Length 405;
 Best Local Similarity 75.0%; Pred. No. 2e+02;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 STEVEGAG 8
 Db 255 STRVEGSG 262
 Search completed: December 11, 2003, 18:30:39
 Job time : 10.2222 secs

;; TITLE OF INVENTION: BACILLUS THURINGIENSIS cryET4 AND cryET5
;; TITLE OF INVENTION: TOXIN GENES AND PROTEINS TOXIC TO LEPIDOPTERAN INSECTS
;; NUMBER OF SEQUENCES: 5
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Panitch Schwarze Jacobs & Nadel c/o A.S.
;; ADDRESSEE: Nadel
;; STREET: 1601 Market Street, 36th Floor
;; CITY: Philadelphia
;; STATE: Pennsylvania
;; COUNTRY: U.S.A.
;; ZIP: 19103
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION NUMBER: US/08/881,340
;; FILING DATE: 24-JUN-1997
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/100,709
;; FILING DATE: 29-JUL-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Egolf, Christopher
;; REGISTRATION NUMBER: 27633
;; REFERENCE/DOCKET NUMBER: 7205-49
;; TELEPHONE: 215-757-1590
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1167 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-881-340-2

Query Match 72.7%; Score 32; DB 2; Length 1167;
Best Local Similarity 75.0%; Pred. No. 3.8e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 TEVEGAGR 9
|||:|
DB 97 TELEGGR 104

RESULT 9
US-08-870-518-9
; Sequence 9, Application US/08870518
; Patent No. 5925566
; GENERAL INFORMATION:
; APPLICANT: Davis, Roger J.
; APPLICANT: Galcheva-Gargova, Zoya
; TITLE OF INVENTION: NON-ACTIVATED RECEPTOR COMPLEX
; TITLE OF INVENTION: PROTEINS AND USES THEREOF
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/870,518
; FILING DATE: 06-JUN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/019,219

;; FILING DATE: 06-JUN-1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Fasse, Peter J.
;; REGISTRATION NUMBER: 32,983
;; REFERENCE/DOCKET NUMBER: 04020/102001
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 617/542-5070
;; TELEFAX: 617/542-8906
;; TELEX: 200154
;; INFORMATION FOR SEQ ID NO: 9:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 207 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-870-518-9

Query Match 70.5%; Score 31; DB 2; Length 207;
Best Local Similarity 55.6%; Pred. No. 96;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 STEVEGAGR 9
:|:|:|
DB 45 NTEIQSAGR 53

RESULT 10
US-09-291-170A-6
; Sequence 6, Application US/09291170A
; Patent No. 6410687
; GENERAL INFORMATION:
; APPLICANT: Vale, Ronald D.
; APPLICANT: Hartman, James J.
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Assays for the Detection of Microtubule
; TITLE OF INVENTION: Depolymerization Inhibitors
; FILE REFERENCE: 18557B-000510US
; CURRENT APPLICATION NUMBER: US/09/291,170A
; CURRENT FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 60/081,734
; PRIOR FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 215
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; OTHER INFORMATION: AAA ATPase superfamily sug1 AAA domain
; US-09-291-170A-6

Query Match 70.5%; Score 31; DB 4; Length 215;
Best Local Similarity 75.0%; Pred. No. 1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 STEVEGAG 8
|||:|
DB 111 STRVEGSG 118

RESULT 11
US-09-724-884-6
; Sequence 6, Application US/09724884
; Patent No. 6429304
; GENERAL INFORMATION:
; APPLICANT: Vale, Ronald D.
; APPLICANT: Hartman, James J.
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Assays for the Detection of Microtubule
; TITLE OF INVENTION: Depolymerization Inhibitors
; FILE REFERENCE: 18557B-000510US
; CURRENT APPLICATION NUMBER: US/09/724,884
; CURRENT FILING DATE: 2000-11-28

TITLE OF INVENTION: BACILLUS THURINGIENSIS cryET4 AND cryET5
 TITLE OF INVENTION: TOXIN GENES AND PROTEINS TOXIC TO LEPIDOPTERAN INSECTS
 NUMBER OF SEQUENCES: 5
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Panitch Schwarze Jacobs & Nadel c/o A.S.
 ADDRESS: Nadel
 STREET: 1601 Market Street, 36th Floor
 CITY: Philadelphia
 STATE: Pennsylvania
 COUNTRY: U.S.A.
 ZIP: 19103
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/100,709
 FILING DATE: 19930729
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Egoif, Christopher
 REGISTRATION NUMBER: 27633
 REFERENCE/DOCKET NUMBER: 7205-49
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 215-757-1590
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1167 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-100-709-2

Query Match 72.7%; Score 32; DB 1; Length 1167;
 Best Local Similarity 75.0%; Pred. No. 3.8e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Qy 2 TEVEGAG 9
 Db 97 TELEGLGR 104
 RESULT 5
 US-08-176-865-2
 ; Sequence 2, Application US/08176865
 ; Patent No. 5616319
 ; GENERAL INFORMATION:
 ; APPLICANT: Donovan, William P.
 ; APPLICANT: Tan, Yuping
 ; APPLICANT: Jany, Christine S.
 ; APPLICANT: Gonzalez Jr., Jose M.
 ; TITLE OF INVENTION: BACILLUS THURINGIENSIS cryET4 AND cryET5
 ; TITLE OF INVENTION: TOXIN GENES AND PROTEINS TOXIC TO LEPIDOPTERAN INSECTS
 ; NUMBER OF SEQUENCES: 5
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Panitch Schwarze Jacobs & Nadel c/o A.S.
 ; ADDRESS: Nadel
 ; STREET: 1601 Market Street, 36th Floor
 ; CITY: Philadelphia
 ; STATE: Pennsylvania
 ; COUNTRY: U.S.A.
 ; ZIP: 19103
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/176,865
 ; FILING DATE: 30-DEC-1993
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:

Query Match 72.7%; Score 32; DB 1; Length 1167;
 Best Local Similarity 75.0%; Pred. No. 3.8e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Qy 2 TEVEGAG 9
 Db 97 TELEGLGR 104
 RESULT 5
 US-08-176-865-2
 ; Sequence 2, Application US/08176865
 ; Patent No. 5616319
 ; GENERAL INFORMATION:
 ; APPLICANT: Donovan, William P.
 ; APPLICANT: Tan, Yuping
 ; APPLICANT: Jany, Christine S.
 ; APPLICANT: Gonzalez Jr., Jose M.
 ; TITLE OF INVENTION: BACILLUS THURINGIENSIS cryET4 AND cryET5
 ; TITLE OF INVENTION: TOXIN GENES AND PROTEINS TOXIC TO LEPIDOPTERAN INSECTS
 ; NUMBER OF SEQUENCES: 5
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Panitch Schwarze Jacobs & Nadel c/o A.S.
 ; ADDRESS: Nadel
 ; STREET: 1601 Market Street, 36th Floor
 ; CITY: Philadelphia
 ; STATE: Pennsylvania
 ; COUNTRY: U.S.A.
 ; ZIP: 19103
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/176,865
 ; FILING DATE: 30-DEC-1993
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:

Query Match 72.7%; Score 32; DB 1; Length 1167;
 Best Local Similarity 75.0%; Pred. No. 3.8e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Qy 2 TEVEGAG 9
 Db 97 TELEGLGR 104
 RESULT 5
 US-08-176-865-2
 ; Sequence 2, Application US/08176865
 ; Patent No. 5616319
 ; GENERAL INFORMATION:
 ; APPLICANT: Donovan, William P.
 ; APPLICANT: Tan, Yuping
 ; APPLICANT: Jany, Christine S.
 ; APPLICANT: Gonzalez Jr., Jose M.
 ; TITLE OF INVENTION: BACILLUS THURINGIENSIS cryET4 AND cryET5
 ; TITLE OF INVENTION: TOXIN GENES AND PROTEINS TOXIC TO LEPIDOPTERAN INSECTS
 ; NUMBER OF SEQUENCES: 5
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Panitch Schwarze Jacobs & Nadel c/o A.S.
 ; ADDRESS: Nadel
 ; STREET: 1601 Market Street, 36th Floor
 ; CITY: Philadelphia
 ; STATE: Pennsylvania
 ; COUNTRY: U.S.A.
 ; ZIP: 19103
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/176,865
 ; FILING DATE: 30-DEC-1993
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:

Query Match 72.7%; Score 32; DB 1; Length 1167;
 Best Local Similarity 75.0%; Pred. No. 3.8e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Qy 2 TEVEGAG 9
 Db 97 TELEGLGR 104
 RESULT 5
 US-08-176-865-2
 ; Sequence 2, Application US/08176865
 ; Patent No. 5616319
 ; GENERAL INFORMATION:
 ; APPLICANT: Donovan, William P.
 ; APPLICANT: Tan, Yuping
 ; APPLICANT: Jany, Christine S.
 ; APPLICANT: Gonzalez Jr., Jose M.
 ; TITLE OF INVENTION: BACILLUS THURINGIENSIS cryET4 AND cryET5
 ; TITLE OF INVENTION: TOXIN GENES AND PROTEINS TOXIC TO LEPIDOPTERAN INSECTS
 ; NUMBER OF SEQUENCES: 5
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Panitch Schwarze Jacobs & Nadel c/o A.S.
 ; ADDRESS: Nadel
 ; STREET: 1601 Market Street, 36th Floor
 ; CITY: Philadelphia
 ; STATE: Pennsylvania
 ; COUNTRY: U.S.A.
 ; ZIP: 19103
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/176,865
 ; FILING DATE: 30-DEC-1993
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:

Query Match 72.7%; Score 32; DB 1; Length 1167;
 Best Local Similarity 75.0%; Pred. No. 3.8e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Qy 2 TEVEGAG 9
 Db 97 TELEGLGR 104
 RESULT 5
 US-08-176-865-2
 ; Sequence 2, Application US/08176865
 ; Patent No. 5616319
 ; GENERAL INFORMATION:
 ; APPLICANT: Donovan, William P.
 ; APPLICANT: Tan, Yuping
 ; APPLICANT: Jany, Christine S.
 ; APPLICANT: Gonzalez Jr., Jose M.
 ; TITLE OF INVENTION: BACILLUS THURINGIENSIS cryET4 AND cryET5
 ; TITLE OF INVENTION: TOXIN GENES AND PROTEINS TOXIC TO LEPIDOPTERAN INSECTS
 ; NUMBER OF SEQUENCES: 5
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Panitch Schwarze Jacobs & Nadel c/o A.S.
 ; ADDRESS: Nadel
 ; STREET: 1601 Market Street, 36th Floor
 ; CITY: Philadelphia
 ; STATE: Pennsylvania
 ; COUNTRY: U.S.A.
 ; ZIP: 19103
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/176,865
 ; FILING DATE: 30-DEC-1993
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:

Query Match 81.8%; Score 36; DB 4; Length 399;
 Best Local Similarity 87.5%; Pred. No. 20;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 STEVEGAG 8
 Db 19 STELEGAG 26
 RESULT 3
 US-08-796-899-25
 ; Sequence 25, Application US/08796899
 ; Patent No. 6160202
 ; GENERAL INFORMATION:
 ; APPLICANT: BUSTOS, Mauricio M
 ; APPLICANT: CHERN, Maw-Sheng
 ; TITLE OF INVENTION: MODIFICATION OF SEED CROPS WITH
 ; TITLE OF INVENTION: TRANSCRIPTION FACTORS
 ; NUMBER OF SEQUENCES: 32
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Burns, Doane, Swecker & Mathis
 ; STREET: P.O. Box 1404
 ; CITY: Alexandria
 ; STATE: Virginia
 ; COUNTRY: United States
 ; ZIP: 22313-1404
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/796,899
 ; FILING DATE: 06-FEB-1997
 ; CLASSIFICATION: 800
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/319,544
 ; FILING DATE: 07-OCT-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Meuth, Donna M
 ; REGISTRATION NUMBER: 36,607
 ; REFERENCE/DOCKET NUMBER: 028754-005
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (703) 836-6620
 ; TELEFAX: (703) 836-2021
 ; INFORMATION FOR SEQ ID NO: 25:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 514 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 US-08-796-899-25

Query Match 75.0%; Score 33; DB 3; Length 514;
 Best Local Similarity 75.0%; Pred. No. 1e-02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 STEVEGAG 8
 Db 266 STELEGK 273
 RESULT 4
 US-08-100-709-2
 ; Sequence 2, Application US/08100709
 ; Patent No. 5322687
 ; GENERAL INFORMATION:
 ; APPLICANT: Donovan, William P.
 ; APPLICANT: Tan, Yuping
 ; APPLICANT: Jany, Christine S.
 ; APPLICANT: Gonzalez Jr., Jose M.
 ; TITLE OF INVENTION: BACILLUS THURINGIENSIS cryET4 AND cryET5
 ; TITLE OF INVENTION: TOXIN GENES AND PROTEINS TOXIC TO LEPIDOPTERAN INSECTS
 ; NUMBER OF SEQUENCES: 5
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Panitch Schwarze Jacobs & Nadel c/o A.S.
 ; ADDRESS: Nadel
 ; STREET: 1601 Market Street, 36th Floor
 ; CITY: Philadelphia
 ; STATE: Pennsylvania
 ; COUNTRY: U.S.A.
 ; ZIP: 19103
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/100,709
 ; FILING DATE: 19930729
 ; CLASSIFICATION: 514
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Egoif, Christopher
 ; REGISTRATION NUMBER: 27633
 ; REFERENCE/DOCKET NUMBER: 7205-49
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 215-757-1590
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1167 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 US-08-100-709-2

Query Match 75.0%; Score 33; DB 3; Length 514;
 Best Local Similarity 75.0%; Pred. No. 1e-02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 STEVEGAG 8
 Db 266 STELEGK 273
 RESULT 4
 US-08-100-709-2
 ; Sequence 2, Application US/08100709
 ; Patent No. 5322687
 ; GENERAL INFORMATION:
 ; APPLICANT: Donovan, William P.
 ; APPLICANT: Tan, Yuping
 ; APPLICANT: Jany, Christine S.
 ; APPLICANT: Gonzalez Jr., Jose M.
 ; TITLE OF INVENTION: BACILLUS THURINGIENSIS cryET4 AND cryET5
 ; TITLE OF INVENTION: TOXIN GENES AND PROTEINS TOXIC TO LEPIDOPTERAN INSECTS
 ; NUMBER OF SEQUENCES: 5
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Panitch Schwarze Jacobs & Nadel c/o A.S.
 ; ADDRESS: Nadel
 ; STREET: 1601 Market Street, 36th Floor
 ; CITY: Philadelphia
 ; STATE: Pennsylvania
 ; COUNTRY: U.S.A.
 ; ZIP: 19103
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/100,709
 ; FILING DATE: 19930729
 ; CLASSIFICATION: 514
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Egoif, Christopher
 ; REGISTRATION NUMBER: 27633
 ; REFERENCE/DOCKET NUMBER: 7205-49
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 215-757-1590
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1167 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 US-08-100-709-2

Query Match 75.0%; Score 33; DB 3; Length 514;
 Best Local Similarity 75.0%; Pred. No. 1e-02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 STEVEGAG 8
 Db 266 STELEGK 273
 RESULT 4
 US-08-100-709-2
 ; Sequence 2, Application US/08100709
 ; Patent No. 5322687
 ; GENERAL INFORMATION:
 ; APPLICANT: Donovan, William P.
 ; APPLICANT: Tan, Yuping
 ; APPLICANT: Jany, Christine S.
 ; APPLICANT: Gonzalez Jr., Jose M.
 ; TITLE OF INVENTION: BACILLUS THURINGIENSIS cryET4 AND cryET5
 ; TITLE OF INVENTION: TOXIN GENES AND PROTEINS TOXIC TO LEPIDOPTERAN INSECTS
 ; NUMBER OF SEQUENCES: 5
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Panitch Schwarze Jacobs & Nadel c/o A.S.
 ; ADDRESS: Nadel
 ; STREET: 1601 Market Street, 36th Floor
 ; CITY: Philadelphia
 ; STATE: Pennsylvania
 ; COUNTRY: U.S.A.
 ; ZIP: 19103
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/100,709
 ; FILING DATE: 19930729
 ; CLASSIFICATION: 514
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Egoif, Christopher
 ; REGISTRATION NUMBER: 27633
 ; REFERENCE/DOCKET NUMBER: 7205-49
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 215-757-1590
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1167 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 US-08-100-709-2

Query Match 75.0%; Score 33; DB 3; Length 514;
 Best Local Similarity 75.0%; Pred. No. 1e-02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 STEVEGAG 8
 Db 266 STELEGK 273
 RESULT 4
 US-08-100-709-2
 ; Sequence 2, Application US/08100709
 ; Patent No. 5322687
 ; GENERAL INFORMATION:
 ; APPLICANT: Donovan, William P.
 ; APPLICANT: Tan, Yuping
 ; APPLICANT: Jany, Christine S.
 ; APPLICANT: Gonzalez Jr., Jose M.
 ; TITLE OF INVENTION: BACILLUS THURINGIENSIS cryET4 AND cryET5
 ; TITLE OF INVENTION: TOXIN GENES AND PROTEINS TOXIC TO LEPIDOPTERAN INSECTS
 ; NUMBER OF SEQUENCES: 5
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Panitch Schwarze Jacobs & Nadel c/o A.S.
 ; ADDRESS: Nadel
 ; STREET: 1601 Market Street, 36th Floor
 ; CITY: Philadelphia
 ; STATE: Pennsylvania
 ; COUNTRY: U.S.A.
 ; ZIP: 19103
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/100,709
 ; FILING DATE: 19930729
 ; CLASSIFICATION: 514
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Egoif, Christopher
 ; REGISTRATION NUMBER: 27633
 ; REFERENCE/DOCKET NUMBER: 7205-49
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 215-757-1590
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1167 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 US-08-100-709-2

Query Match 75.0%; Score 33; DB 3; Length 514;
 Best Local Similarity 75.0%; Pred. No. 1e-02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 STEVEGAG 8
 Db 266 STELEGK 273
 RESULT 4
 US-08-100-709-2
 ; Sequence 2, Application US/08100709
 ; Patent No. 5322687
 ; GENERAL INFORMATION:
 ; APPLICANT: Donovan, William P.
 ; APPLICANT: Tan, Yuping
 ; APPLICANT: Jany, Christine S.
 ; APPLICANT: Gonzalez Jr., Jose M.
 ; TITLE OF INVENTION: BACILLUS THURINGIENSIS cryET4 AND cryET5
 ; TITLE OF INVENTION: TOXIN GENES AND PROTEINS TOXIC TO LEPIDOPTERAN INSECTS
 ; NUMBER OF SEQUENCES: 5
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Panitch Schwarze Jacobs & Nadel c/o A.S.
 ; ADDRESS: Nadel
 ; STREET: 1601 Market Street, 36th Floor
 ; CITY: Philadelphia
 ; STATE: Pennsylvania
 ; COUNTRY: U.S.A.
 ; ZIP: 19103
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/100,709
 ; FILING DATE: 19930729
 ; CLASSIFICATION: 514
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Egoif, Christopher
 ; REGISTRATION NUMBER: 27633
 ; REFERENCE/DOCKET NUMBER: 7205-49
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 215-757-1590
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1167 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 US-08-100-709-2

CC the amino acid sequence of bZIP1 (basic leucine zipper) ORF3 protein.
 XX Sequence 514 AA;
 SQ

Query Match 75.0%; Score 33; DB 22; Length 514;
 Best Local Similarity 75.0%; Pred. No. 3.7e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 STEVEGAG 8
 |||:|:
 Db 266 STEIEGKG 273

RESULT 14
 ID ABB71150
 XX ABB71150 standard; Protein; 2858 AA.
 AC ABB71150;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 40242.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI; 2001-656860/75.
 DR N-PSDB; ABL15253.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Disclosure; SEQ ID NO 40242; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 PS Disclosure; SEQ ID NO 40242; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 2858 AA;

Query Match 75.0%; Score 33; DB 22; Length 2858;
 Best Local Similarity 55.6%; Pred. No. 2.3e+03;
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 STEVEGAG 9
 |||:|:
 Db 1309 TTEIEGSGQ 1317

RESULT 15
 ID ABB58064
 XX ABB58064 standard; Protein; 3060 AA.
 AC ABB58064;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 984.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI; 2001-656860/75.
 DR N-PSDB; ABL02167.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Disclosure; SEQ ID NO 984; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 3060 AA;
 Query Match 75.0%; Score 33; DB 22; Length 3060;
 Best Local Similarity 55.6%; Pred. No. 2.5e+03;
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 STEVEGAG 9
 |||:|:
 Db 1309 TTEIEGSGQ 1317

Search completed: December 11, 2003, 18:24:58
 Job time : 26 secs

PN WO958098-A2.
 XX 18-NOV-1999.
 XX 12-MAY-1999; 99WO-US10549.
 XX 12-MAY-1999; 98US-0085197.
 PR 05-MAY-1999; 99US-0085197.
 XX (BOCK/) BOCK S C.
 PA (PICA/) PICARD V.
 PA (ZEND/) ZENDEHROUH P.
 XX Bock SC, Picard V, Zendeirouh P;
 PI WPI; 2000-116274/10.
 DR New modified human antithrombin III compounds, used for treating e.g.
 XX sepsis, trauma, acute respiratory distress syndrome, restenosis,
 PT thrombosis, thromboembolism or stroke -
 XX Claim 13; Page 57; 75pp; English.
 XX The present sequence is from an antithrombin III (ATIII) variant, Bb.B
 CC derived from human ATIII.N135A cDNA insert of the pBlueBac baculovirus
 CC expression construct and comprises residues 385-395. The variant has
 CC improved resistance to elastase and IgG-activated neutrophils while
 CC retains anti-thrombin and anti-factor Xa activities. It may be
 CC expressed as glycoforms with enhanced heparin affinity which target the
 CC blood vessel wall more efficiently than ATIIIs with normal heparin
 CC affinity. The modified ATIIIs can be used to treat thrombin activation-
 CC related pathological symptoms due to sepsis, trauma, acute
 CC respiratory distress syndrome, restenosis, thrombosis, thromboembolism
 CC and stroke. It can also be used to reduce the risk of reocclusion
 CC and restenosis in percutaneous transluminal coronary angioplasty,
 CC thrombosis associated with surgery, ischaemia/reperfusion injury, and
 CC coagulation abnormalities in cancer or surgical patients.
 XX Sequence 9 AA;
 SQ Query Match 77.3%; Score 34; DB 21; Length 9;
 Best Local Similarity 77.8%; Pred. No. 9.3e+05;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 STEVEGAGR 9
 DB 1 STEGEGSGR 9
 RESULT 10
 AAG90727
 ID AAG90727 standard; Protein; 135 AA.
 XX AAG90727;
 AC AAG90727;
 XX 26-SEP-2001 (first entry)
 DT C glutamicum protein fragment SEQ ID NO: 4481.
 DE Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
 XX organic acid synthesis.
 KW Coryneform bacterium glutamicum.
 XX EP1108790-A2.
 XX 20-JUN-2001.
 PD 18-DEC-2000; 2000EP-0127688.
 XX 16-DEC-1999; 99JP-0377484.
 PR 07-APR-2000; 2000JP-0159162.
 PR 03-AUG-2000; 2000JP-0280988.
 XX

(KYOW) KYOWA HAKKO KOGYO KK.
 XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A,
 XX WPI; 2001-376931/40.
 DR N-PSDB; AAH65946.
 XX Novel polynucleotides derived from Coryneform bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analysing
 PT expression profile of pattern of a gene and identifying homologous gene
 PT Claim 17; SEQ ID NO: 4481; 246pp + Sequence Listing; English.
 XX The present invention provides a number of nucleotide and protein
 CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of coryneform bacterium, measuring expression amount and
 CC analysing the expression profile or expression pattern of a gene derived
 CC from Coryneform bacterium, and identifying a homologue of a gene derived
 CC from coryneform bacterium. Coryneform bacteria are useful for producing
 CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a protein described
 CC in the exemplification of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the
 CC European Patent Office.
 XX Sequence 135 AA;
 SQ Query Match 77.3%; Score 34; DB 22; Length 135;
 Best Local Similarity 87.5%; Pred. No. 55;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 STEVEGAG 8
 DB 48 STEVEGAG 55
 RESULT 11
 ABB98317
 ID ABB98317 standard; Protein; 730 AA.
 XX ABB98317;
 AC ABB98317;
 XX 09-JAN-2003 (first entry)
 DT Human leukocyte derived protein SEQ ID NO 8768.
 XX Human; leukocyte; gene expression profiling; allograft rejection;
 DE atherosclerosis; congestive heart failure; systemic lupus erythematosus;
 KW rheumatoid arthritis; osteoarthritis; cytomegalovirus; infection.
 XX Homo sapiens.
 OS WO200257414-A2.
 XX 25-JUL-2002.
 PD 22-OCT-2001; 2001WO-US47856.
 XX 20-OCT-2000; 2000US-241994P.
 PR 08-JUN-2001; 2001US-296764P.
 XX (BIOC-) BIOCARDIA INC.
 XX Wohlgenuth J, Fry K, Matcuk G, Altman P, Prentice J, Phillips J;
 PI Ly N, Woodward R, Quattermous T, Johnson F;
 XX WPI; 2002-636525/68.
 XX


```

PD 27-JAN-2000.
XX
XX PF 15-JUL-1999; 99WO-US15974.
XX
XX PR 16-JUL-1998; 98US-0092980.
XX
XX PA (SLOK ) SLOAN KETTERING INST CANCER RES.
XX
XX PI Pasternak G, Pan Y;
XX
XX DR WPI; 2000-182402/16.
XX
XX DR N-PSDB; AAZ60730.
XX
XX PT New splice variants of the mu-opioid receptor, useful in screening for
XX PT selective analgesics and for regulating morphine analgesia or body
XX PT weight -
XX
XX PS Example 1; Fig 3A; 83pp; English.
XX
XX CC The present sequence represents a murine mu-opioid receptor (MOR-1)
XX CC splice variant MOR-1G. The specification describes 11 new exons for
XX CC the MOR-1 gene, which combine to yield 15 novel splice variants of
XX CC the MOR-1 gene. These splice variants are potential targets for
XX CC modulating morphine analgesia and opioid-mediated ingestive responses.
XX CC The MOR-1 polypeptide is used to screen compounds for opioid activity.
XX CC Such compounds are potential analgesics or more generally agents that
XX CC affect gastrointestinal motility, respiration or the immune, endocrine
XX CC or autonomic nervous systems, e.g. regulators of peristalsis.
XX CC Antagonists, agonists and ligands of MOR-1, as well as DNA vectors
XX CC expressing MOR-1-encoding nucleic acids, or sequences antisense to
XX CC MOR-1 nucleic acids, are used to regulate morphine analgesia and body
XX CC weight. The level of MOR-1 or tissue distribution of MOR-1 can be
XX CC measured to diagnose MOR-1 related pharmacological abnormalities or
XX CC neuroendocrine disorders, particularly inherited disorders. Transgenic
XX CC animals with extra copies of the MOR-1 gene, or with endogenous alleles
XX CC deleted, are used to study loss or gain of function phenotypes.
XX
XX SQ Sequence 359 AA;
XX
XX Query Match 81.8%; Score 36; DB 21; Length 359;
XX Best Local Similarity 87.5%; Pred. No. 63;
XX Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 STEVEGAG 8
XX |||:||||
XX 19 STELEGAG 26
XX
XX RESULT 6
XX AAY68881
XX ID AAY68881 standard; Protein; 399 AA.
XX
XX AC AAY68881;
XX
XX DT 16-MAY-2000 (first entry)
XX
XX DE A murine mu-opioid receptor splice variant MOR-1H.
XX
XX KW Mu-opioid receptor; MOR-1; splice variant; morphine analgesia;
XX KW opioid-mediated ingestive response; opioid activity; analgesic;
XX KW gastrointestinal motility; respiration; immune system;
XX KW endocrine system; autonomic nervous system; peristalsis regulator;
XX KW body weight; neuroendocrine disorder; MOR-1H.
XX
XX OS Mus sp.
XX
XX FH Key Location/Qualifiers
XX FT Misc-difference 390
XX FT /note= "encoded by GAA"
XX
XX PN WO200004046-A2.
XX
XX PD 27-JAN-2000.
XX
XX PF 15-JUL-1999; 99WO-US15974.
XX
XX PR 16-JUL-1998; 98US-0092980.
XX
XX PA (SLOK ) SLOAN KETTERING INST CANCER RES.
XX
XX PI Pasternak G, Pan Y;
XX
XX DR WPI; 2000-182402/16.
XX
XX DR N-PSDB; AAZ60730.
XX
XX PT New splice variants of the mu-opioid receptor, useful in screening for
XX PT selective analgesics and for regulating morphine analgesia or body
XX PT weight -
XX
XX PS Example 1; Fig 3A; 83pp; English.
XX
XX CC The present sequence represents a murine mu-opioid receptor (MOR-1)
XX CC splice variant MOR-1G. The specification describes 11 new exons for
XX CC the MOR-1 gene, which combine to yield 15 novel splice variants of
XX CC the MOR-1 gene. These splice variants are potential targets for
XX CC modulating morphine analgesia and opioid-mediated ingestive responses.
XX CC The MOR-1 polypeptide is used to screen compounds for opioid activity.
XX CC Such compounds are potential analgesics or more generally agents that
XX CC affect gastrointestinal motility, respiration or the immune, endocrine
XX CC or autonomic nervous systems, e.g. regulators of peristalsis.
XX CC Antagonists, agonists and ligands of MOR-1, as well as DNA vectors
XX CC expressing MOR-1-encoding nucleic acids, or sequences antisense to
XX CC MOR-1 nucleic acids, are used to regulate morphine analgesia and body
XX CC weight. The level of MOR-1 or tissue distribution of MOR-1 can be
XX CC measured to diagnose MOR-1 related pharmacological abnormalities or
XX CC neuroendocrine disorders, particularly inherited disorders. Transgenic
XX CC animals with extra copies of the MOR-1 gene, or with endogenous alleles
XX CC deleted, are used to study loss or gain of function phenotypes.
XX
XX SQ Sequence 399 AA;
XX
XX Query Match 81.8%; Score 36; DB 21; Length 399;
XX Best Local Similarity 87.5%; Pred. No. 71;
XX Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 STEVEGAG 8
XX |||:||||
XX 19 STELEGAG 26
XX
XX RESULT 7
XX AAY44476
XX ID AAY44476 standard; peptide; 9 AA.
XX
XX AC AAY44476;
XX
XX DT 27-MAR-2000 (first entry)
XX
XX DE Human antithrombin III variant 13.D (residues 385-393).
XX
XX KW Human; antithrombin III; ATIII variant 13.D; elastase-resistant;
XX KW IGC activated neutrophil resistant; anti-thrombin activity; heparin;
XX KW anti-factor Xa activity; blood clotting disorder; sepsis; trauma; stroke;
XX KW thrombin activation-related pathological symptom; restenosis; thrombosis;
XX KW acute respiratory distress syndrome; thromboembolism; reocclusion.
XX
XX OS Homo sapiens.
XX
XX OS Synthetic.
XX
XX FH Key Location/Qualifiers
XX FT Misc-difference 4
XX FT /note= "ATIII.N135A Val at 388 is substituted by Leu"
XX
XX FT Misc-difference 5
XX FT /note= "ATIII.N135A Val at 389 is substituted by Glu"
XX
XX FT Misc-difference 6
XX FT /note= "ATIII.N135A Ile at 390 is substituted by Gly"
XX
XX PD 27-JAN-2000.

```

XX 12-MAY-1998; 98US-0085197.
 PR 05-MAY-1999; 99US-0085197.
 XX (BOCK/) BOCK S C.
 PA (PICA/) PICARD V.
 PA (ZEND/) ZENDEHROUH P.
 XX Bock SC, Picard V, Zendeherouh P;
 PI WPI; 2000-116274/10.
 DR
 XX New modified human antithrombin III compounds, used for treating e.g.
 PT sepsis, trauma, acute respiratory distress syndrome, restenosis,
 PT thrombosis, thromboembolism or stroke -
 XX
 PS Claim 13; Page 57; 75pp; English.
 XX The present sequence is from an antithrombin III (ATIII) variant, 13.B
 CC derived from human ATIII.N135A cDNA insert of the pBlueBac baculovirus
 CC expression construct and comprises residues 385-393. The variant has
 CC improved resistance to elastase and Igg-activated neutrophils while
 CC retains anti-thrombin and anti-factor Xa activities. It may be
 CC expressed as glycoforms with enhanced heparin affinity which target the
 CC blood vessel wall more efficiently than ATIII with normal heparin
 CC affinity. The modified ATIII can be used to treat thrombin activation-
 CC related pathological symptoms due to sepsis, trauma, acute
 CC respiratory distress syndrome, restenosis, thrombosis, thromboembolism
 CC and stroke. It can also be used to reduce the risk of reocclusion
 CC and restenosis in percutaneous transluminal coronary angioplasty,
 CC thrombosis associated with surgery, ischaemia/reperfusion injury, and
 CC coagulation abnormalities in cancer or surgical patients.
 XX
 SQ Sequence 9 AA;
 Query Match 100.0%; Score 44; DB 21; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 STEVEGAGR 9
 Db |||||
 1 STEVEGAGR 9
 RESULT 2
 AAY44475
 ID AAY44475 standard; peptide; 9 AA.
 XX
 AC AAY44475;
 XX
 DT 27-MAR-2000 (first entry)
 XX
 DE Human antithrombin III variant 13.C (residues 385-393).
 XX
 KW Human; antithrombin III; ATIII variant 13.C; elastase-resistant;
 KW Igg activated neutrophil resistant; anti-thrombin activity; heparin;
 KW anti-factor Xa activity; blood clotting disorder; sepsis; trauma; stroke;
 KW thrombin activation-related pathological symptom; restenosis; thrombosis;
 KW acute respiratory distress syndrome; thromboembolism; reocclusion.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 3 /note= "ATIII.N135A Ala at 387 is substituted by Glu"
 FT Misc-difference 4 /note= "ATIII.N135A Val at 388 is substituted by Leu"
 FT Misc-difference 5 /note= "ATIII.N135A Val at 389 is substituted by Glu"
 FT Misc-difference 6 /note= "ATIII.N135A Ile at 390 is substituted by Gly"
 XX

PN WO9958098-A2.
 XX
 PD 18-NOV-1999.
 XX
 PF 12-MAY-1999; 99WO-US10549.
 XX
 PR 12-MAY-1998; 98US-0085197.
 PR 05-MAY-1999; 99US-0085197.
 XX
 XX (BOCK/) BOCK S C.
 PA (PICA/) PICARD V.
 PA (ZEND/) ZENDEHROUH P.
 XX Bock SC, Picard V, Zendeherouh P;
 PI WPI; 2000-116274/10.
 DR
 XX New modified human antithrombin III compounds, used for treating e.g.
 PT sepsis, trauma, acute respiratory distress syndrome, restenosis,
 PT thrombosis, thromboembolism or stroke -
 XX
 PS Claim 13; Page 57; 75pp; English.
 XX The present sequence is from an antithrombin III (ATIII) variant, 13.C
 CC derived from human ATIII.N135A cDNA insert of the pBlueBac baculovirus
 CC expression construct and comprises residues 385-393. The variant has
 CC improved resistance to elastase and Igg-activated neutrophils while
 CC retains anti-thrombin and anti-factor Xa activities. It may be
 CC expressed as glycoforms with enhanced heparin affinity which target the
 CC blood vessel wall more efficiently than ATIII with normal heparin
 CC affinity. The modified ATIII can be used to treat thrombin activation-
 CC related pathological symptoms due to sepsis, trauma, acute
 CC respiratory distress syndrome, restenosis, thrombosis, thromboembolism
 CC and stroke. It can also be used to reduce the risk of reocclusion
 CC and restenosis in percutaneous transluminal coronary angioplasty,
 CC thrombosis associated with surgery, ischaemia/reperfusion injury, and
 CC coagulation abnormalities in cancer or surgical patients.
 XX
 SQ Sequence 9 AA;
 Query Match 93.2%; Score 41; DB 21; Length 9;
 Best Local Similarity 88.9%; Pred. No. 9.3e+05;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 STEVEGAGR 9
 Db |||||
 1 STEVEGAGR 9
 RESULT 3
 AAY44466
 ID AAY44466 standard; peptide; 9 AA.
 XX
 AC AAY44466;
 XX
 DT 27-MAR-2000 (first entry)
 XX
 DE Human antithrombin III variant 7EVEA (residues 385-393).
 XX
 KW Human; antithrombin III; ATIII variant 7EVEA; elastase-resistant;
 KW Igg activated neutrophil resistant; anti-thrombin activity; heparin;
 KW anti-factor Xa activity; blood clotting disorder; sepsis; trauma; stroke;
 KW thrombin activation-related pathological symptom; restenosis; thrombosis;
 KW acute respiratory distress syndrome; thromboembolism; reocclusion.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 3 /note= "ATIII.N135A Ala at 387 is substituted by Glu"
 FT Misc-difference 5 /note= "ATIII.N135A Val at 389 is substituted by Glu"
 FT

DR Pfam; PF00844; Gemini_coat; 1.
 DR PRINTS; PRO0223; GEMCOATARBRI.
 DR PRODOM; PD001964; Gemcoat_MSV; 1.
 SQ SEQUENCE 260 AA; 29420 MW; EFOFE49A99B14197 CRC64;

Query Match 78.3%; Score 36; DB 12; Length 260;
 Best Local Similarity 87.5%; Pred. No. 88;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 EGECSGR 9
 |||||
 Db 16 EGECSGR 23

Search completed: December 11, 2003, 18:28:22
 Job time : 23.3333 secs

DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Putative ripening-related bZIP protein.
GN GRIP5.
OS *Vitis vinifera* (Grape).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Vitaceae;
OC *Vitis*.
OX NCBI_TaxID=29760;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Shiraz; TISSUE=fruit;
RX MEDLINE=20177861; PubMed=10712544;
RA Davies C., Robinson S.P.;
RT differential screening indicates a dramatic change in mRNA profiles
RT during grape berry ripening. Cloning and characterization of cDNAs
RT encoding putative cell wall and stress response proteins.";
RL plant Physiol. 122:803-812(2000).
CC -1- SIMILARITY: BELONGS TO THE BZIP FAMILY.
DR EMBL; AJ237992; CAB85632.1; -.
DR HSP; P03069; 2DGC.
DR InterPro; IPR004827; TF_BZIP.
DR Pfam; PF00170; bZIP; 1.
DR SMART; SM00338; BRIZ; 1.
DR PROSITE; PS00036; BZIP_BASIC; 1.
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 447 AA; 47879 MW; 68A1962DF4BFC08 CRC64;

Query Match 80.4%; Score 37; DB 10; Length 447;
Best Local Similarity 66.7%; Pred. No. 1e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 SEEGSGS 9
Db 13 SQQGDGSGR 21
|:|:|:|:|
|:|:|:|:|

RESULT 9
Q83S1 PRELIMINARY; PRT; 711 AA.
AC Q83S1; 2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DE 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Putative kinesin light chain gene.
GN 49d11.11.
OS *Oryza sativa* (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; *Oryza*.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Park Y.-J., Roskos N., Ramakrishna W., SanMiguel P., Shiloff B.,
RA Ma J., Jiang Z., Kleinhofs A., Bennetzen J.;
RT "Sequence characterization of orthologous regions in the barley and
RT rice genomes."
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF480496; AAL87157.1; -.
DR Gramene; Q83S1; -.
DR InterPro; IPR001440; TPR.
DR Pfam; PF00515; TPR; 9.
DR SMART; SM00028; TPR; 9.
SQ SEQUENCE 711 AA; 76952 MW; 2EBD370C75AA0EB4 CRC64;

Query Match 80.4%; Score 37; DB 10; Length 711;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 BEEGSGS 8
Db 112 BEEGSG 118
|:|:|:|:|
|:|:|:|:|

RESULT 10
Q8H922 PRELIMINARY; PRT; 731 AA.
ID Q8H922
AC Q8H922;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical protein.
GN OSJNBA007K18.2.
OS *Oryza sativa* (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; *Oryza*.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA Overton II L.L., Teitlin T., Kim M.M., Bera J.J., Jin S.S.,
RA Padrosh D.W., Tallon L.J., Koo H., Zismann V., Hsiao J., Blunt S.,
RA Vanaken S.S., Riedmuller S.B., Utterback T.T., Feldblyum T.V.,
RA Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
RA White O., Salzberg S.L., Fraser C.M.;
RT "Oryza sativa chromosome 10 BAC OSJNBA007K18 genomic sequence."
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC027038; AAN05523.1; -.
KW Hypothetical protein.
SQ SEQUENCE 731 AA; 78227 MW; 335F9C377831ADD7 CRC64;

Query Match 80.4%; Score 37; DB 10; Length 731;
Best Local Similarity 87.5%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 BEEGSGS 9
Db 36 BEEGSGR 43
|:|:|:|:|
|:|:|:|:|

RESULT 11
Q9LWH6 PRELIMINARY; PRT; 1023 AA.
ID Q9LWH6
AC Q9LWH6; 2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical protein.
OS *Oryza sativa* (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; *Oryza*.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, PAC
RT clone: P0675A05.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP002071; BAA59883.1; -.
DR Gramene; Q9LWH6; -.
DR InterPro; IPR001810; F-box.
DR InterPro; IPR006527; F-box_assoc_1.
DR Pfam; PF00646; F-box; 1.
DR SMART; SM00256; FBOX; 1.
DR TIGRfam; TIGR01640; F-box_assoc_1; 1.
DR PROSITE; PS0181; FBOX; 1.
KW Hypothetical protein.
SQ SEQUENCE 1023 AA; 112445 MW; FAB453DB4A5ED889 CRC64;

Query Match 80.4%; Score 37; DB 10; Length 1023;
Best Local Similarity 77.8%; Pred. No. 2.5e+02;

```

OS Oryza sativa (Pice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoidae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RR [1]
RN SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RC Sasaki T.; Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
RT clones:P0407B12.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RR EMBL; AP002843; BAB17181.1; -.
DR Gramene; Q9FTW2; -.
SQ SEQUENCE      284 AA; 30171 MW; 33EB3325109A02FC CRC64;

Query Match          84.8%; Score 39; DB 10; Length 284;
Best Local Similarity 77.8%; Pred. No. 27;
Matches       7; Conservative    1; Mismatches     1; Indels   0; Gaps   0;

QY 1 SEQEGSGR 9
DB 254 ABEGECDGR 262
      :|||||
      |||||

RESULT 3
Q40870 PRELIMINARY; PRT; 509 AA.
ID Q40870
AC Q40870;
DC 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Legumin-like storage protein.
DE Picea glauca (White spruce).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Pinetales; Pinaceae; Picea.
OX NCBI_TaxID=3330;
RR [1]
RN SEQUENCE FROM N.A.
RC STRAINS=Pg118; TISSUE=Somatic embryo;
RA Newton C.H.;
RL Submitted (NOV-1991) to the EMBL/GenBank/DBJ databases.
DR EMBL; X63192; CA44874.1; -.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR006044; Seedstore_1ls.
DR Pfam; PF00190; Cupin; 2.
DR PRINTS; PR00439; 11SGOBULIN.
SQ SEQUENCE      509 AA; 57369 MW; C467AE867CA81027 CRC64;

Query Match          82.6%; Score 38; DB 10; Length 509;
Best Local Similarity 87.5%; Pred. No. 77;
Matches       7; Conservative    0; Mismatches     1; Indels   0; Gaps   0;

QY 2 BEGECSGR 9
DB 221 BEGEGCR 228
      |||||
      |||||

RESULT 4
Q8RAM1 PRELIMINARY; PRT; 820 AA.
ID Q8RAM1
AC Q8RAM1;
DC 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Single-stranded DNA-specific exonuclease.
GN RECJ OR TTE1191.
OC Thermoanaerobacter tengcongensis.
OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacterales;
OC Thermoanaerobacteriaceae; Thermoanaerobacter.
OX NCBI_TaxID=119072;
RR [1]
RN SEQUENCE FROM N.A.
```

FT NP_BIND 32 39 ATP (POTENTIAL).
 FT DOMAIN 173 357 COILED COIL (POTENTIAL).
 FT DOMAIN 389 504 COILED COIL (POTENTIAL).
 FT DOMAIN 505 667 FLEXIBLE HINGE.
 FT DOMAIN 668 1022 COILED COIL (POTENTIAL).
 FT DOMAIN 1115 1150 ALA/ASP-RICH (DA-BOX).
 SQ SEQUENCE 1191 AA; 138447 MW; 1F33AEF05E9198A2 CRC64;

Query Match 73.9%; Score 34; DB 1; Length 1191;
 Best Local Similarity 85.7%; Pred. No. 2.2e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 ERGEGSG 8
 Db 1069 DEEGSG 1075

Search completed: December 11, 2003, 18:21:01
 Job time : 6.66667 secs

Hic2, a novel homolog of Hic1 and gammaPBP.";
Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
-1- FUNCTION: Transcriptional repressor (By similarity).
-1- SUBUNIT: Interacts with CtBP (By similarity).
-1- SUBCELLULAR LOCATION: Nuclear (Probable).
-1- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-FINGER PROTEINS. HIC TRANSCRIPTION FACTORS SUBFAMILY.
-1- SIMILARITY: Contains 5 C2H2-type zinc fingers.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AF117382; AAF28801.1; -;
MGD; MG11929869; Hic2.
InterPro; IPR007087; Znf_C2H2.
Pfam; PF00086; zf-C2H2; 5.
SMART; SM00355; Znf_C2H2; 5.
PROSITE; PS00028; ZINC_FINGER_C2H2_1; 5.
PROSITE; PS00157; ZINC_FINGER_C2H2_2; 5.
Multigene family; Nuclear protein; Transcription regulation;
DNA-binding; Zinc-finger; Metal-binding; Repeat; Repressor.
NON_TER 1
1 ZN_FING 93 C2H2-TYPE 1.
2 ZN_FING 120 C2H2-TYPE 2.
3 ZN_FING 156 183 C2H2-TYPE 3.
4 ZN_FING 184 211 C2H2-TYPE 4.
5 ZN_FING 212 239 C2H2-TYPE 5.
6 ZN_FING 240 266 C2H2-TYPE 5.
SEQUENCE 266 AA; 28946 MW; 507B291FAC5149AD CRC64;

Query Match 73.9%; Score 34; DB 1; Length 266;
Best Local Similarity 85.7%; Pred. No. 44;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 EEGRGSG 8
|||||
12 EEGRGTG 18

DDB

RESULT 13
ALDX_SPOSA
ID ALDX_SPOSA STANDARD; PRT; 322 AA.
P27800; Q12707;
01-AUG-1992 (Rel. 23, Created)
28-FEB-2003 (Rel. 41, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
Aldehyde reductase I (EC 1.1.1.2) (Alcohol dehydrogenase [NADP+]) (AUR).
DE DE
GN GN
OS Sporobolomyces salmonicolar.
OC Eukaryota; Fungi; Basidiomycota; Urediniomycetes;
OC Microbotryomycetidae; Microbotryomycetidae incertae sedis;
OC Sporidiobolaceae; Sporidiobolus.
OC NCBI_TaxID=5005;
[1]
SEQUENCE FROM N.A., SEQUENCE OF 33-124; 132-153; 156-199 AND 202-313, AND FUNCTION.
RC STRAIN=AKU 4429;
RC MEDLINE=96271678; PubMed=8779568;
RA Kita K., Matsuzaki K., Hashimoto T., Yanase H., Kato N.,
RA Chung M.C.-M., Kataoka M., Shimizu S.;
RT "Cloning of the aldehyde reductase gene from a red yeast,
RT Sporobolomyces salmonicolar, and characterization of the gene and its
RT product.";
RT Appl. Environ. Microbiol. 62:2303-2310 (1996).
[2]
RN
RN SEQUENCE OF 1-51, AND ENZYME REGULATION.
RP RP
RC STRAIN=AKU 4429;
RC MEDLINE=92338224; PubMed=1633196;
RX

```

FT CARBOHYD 470 470 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 486 486 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 489 489 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 635 635 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 644 644 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 660 660 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 769 769 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 889 AA; 102053 MW; 1438BD23B900D0AA CRC64;

Query Match 76.1%; Score 35; DB 1; Length 889;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 BEBEGSGR 9
DB 759 BEBDSGGR 766

RESULT 10
HAIR HUMAN STANDARD; PRT; 1189 AA.
AC O43593; Q9NP61;
DT 15-JUL-1999 (Rel. 38, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hairless protein.
GN HR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM LONG), AND VARIANT ALUNC ALA-1022.
RX MEDLINE=98111413; PubMed=9445480;
RA Ahmad W., ul Haque M.F., Brancolini V., Tsoi H.C., Ul Haque S.,
RA Lam H., Aita V.M., Owen J., Deblaquiere M., Frank J.,
RA Ceerhalni-Friedman P.B., Leask A., McGrath J.A., Peacocke M.,
RA Ahmad M., Ott J., Christiano A.M.;
RT "Alopecia universalis associated with a mutation in the human hairless
RT gene.";
RL Science 279:720-724(1998).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM LONG), REVISIONS TO 572 AND 774, AND
RP TISSUE SPECIFICITY.
RX TISSUE=Peripheral blood leukocytes, and skin fibroblast;
RX MEDLINE=99152400; PubMed=10051399;
RA Ahmad W., Zlotogorski A., Panteleyev A.A., Lam H., Ahmad M.,
RA ul Haque M.F., Abdallah H.M., Dragan L., Christiano A.M.;
RT "Genomic organization of the human hairless gene (HR) and
RT identification of a mutation underlying congenital atrichia in an Arab
RT Palestinian family.";
RL Genomics 56:141-148(1999).
RN [3]
RP SEQUENCE FROM N.A., VARIANT ALUNC ASP-1136, TISSUE SPECIFICITY, AND
RP ALTERNATIVE SPLICING.
RX TISSUE=Brain, Fetal brain, and Peripheral blood leukocytes;
RX MEDLINE=98409496; PubMed=9736769;
RA Cichon S., Anker M., Vogt I.R., Rohleder H., Putzstuck M., Hillmer A.,
RA Parooq S.A., Al-Dhafri K.S., Ahmad M., Haque S., Rietschel M.,
RA Clopping P., Kruse R., Noethen M.M.;
RT "Cloning, genomic organization, alternative transcripts and mutational
RT analysis of the gene responsible for autosomal recessive universal
RT congenital alopecia.";
RL Hum. Mol. Genet. 7:1671-1679(1998).
RN [4]
RP VARIANT APL GLN-620.
RX MEDLINE=98431781; PubMed=9758627;
RA Ahmad W., Irvine A.D., Lam H., Buckley C., Bingham E.A.,
RA Panteleyev A.A., Ahmad M., McGrath J.A., Christiano A.M.;
RT "A missense mutation in the zinc-finger domain of the human hairless
RT gene underlies congenital atrichia in a family of Irish travellers.";
RL Am. J. Hum. Genet. 63:984-991(1998).
RN [5]
RP -I- FUNCTION: MAY ACT AS A TRANSCRIPTION FACTOR THAT COULD ACT ON TO
CC

```

```

CC REGULATE ONE OF THE PHASES OF HAIR GROWTH.
CC -I- SUBCELLULAR LOCATION: Nuclear.
CC -I- ALTERNATIVE PRODUCTS.
CC Event-Alternative splicing; Named isoforms=2;
CC Comment-Additional isoforms seem to exist;
CC Name=Long;
CC IsoId=O43593-1; Sequence=Displayed;
CC Name=Short;
CC IsoId=O43593-2; Sequence=VSP_004276;
CC -I- TISSUE SPECIFICITY: Strongest expression of both isoforms is seen
CC in the small intestine, weaker expression in brain and colon, and
CC trace expression is found in liver, pancreas, spleen, thymus,
CC stomach, salivary gland, appendix and trachea. Long isoform is
CC always the most abundant. Long isoform is exclusively expressed at
CC low levels in kidney and testis and short isoform exclusively at
CC high levels in the skin.
CC -I- DISEASE: DEFECTS IN HR ARE THE CAUSE OF ALOPECIA UNIVERSALIS
CC (ALUNC). A RARE AUTOSOMAL RECESSIVE FORM OF HAIR LOSS. IT IS
CC CHARACTERIZED BY HAIR FOLLICLES WITHOUT HAIR.
CC -I- DISEASE: DEFECTS IN HR ARE THE CAUSE OF ATRICHIA WITH PAPULAR
CC LESIONS (APL) (ALSO KNOWN AS CONGENITAL ATRICHIA). THIS AUTOSOMAL
CC RECESSIVE DISEASE IS CHARACTERIZED BY PAPILLARY LESIONS OVER MOST
CC OF THE BODY AND ALMOST COMPLETE ABSENCE OF HAIR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF039196; AAC32258.2; -
CC EMBL; AJ277249; CAB87577.2; -
CC EMBL; AJ277250; CAB87577.2; JOINED.
CC EMBL; AJ277251; CAB87577.2; JOINED.
CC EMBL; AJ277252; CAB87577.2; JOINED.
CC EMBL; AJ277253; CAB87577.2; JOINED.
CC EMBL; AJ400825; CAB87577.2; JOINED.
CC EMBL; AJ400826; CAB87577.2; JOINED.
CC EMBL; AJ400827; CAB87577.2; JOINED.
CC EMBL; AJ400828; CAB87577.2; JOINED.
CC EMBL; AJ400829; CAB87577.2; JOINED.
CC EMBL; AJ400830; CAB87577.2; JOINED.
CC EMBL; AJ400831; CAB87577.2; JOINED.
CC EMBL; AJ400832; CAB87577.2; JOINED.
CC EMBL; AJ400833; CAB87577.2; JOINED.
CC EMBL; AJ400834; CAB87577.2; JOINED.
CC EMBL; AJ400835; CAB87577.2; JOINED.
CC EMBL; AJ400836; CAB87577.2; JOINED.
CC EMBL; AJ400837; CAB87577.2; JOINED.
CC EMBL; AJ277165; CAB86602.1; -
CC Genew; HGNC:5172; HR.
CC MIN; 602302; -
CC MIN; 203455; -
CC MIN; 209500; -
CC GO; GO:0005634; C:nucleus; NAS.
CC GO; GO:0003700; F:transcription factor activity; NAS.
CC GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
CC InterPro; IPR003347; TF_JmJc.
CC Pfam; PF02373; JmJc_1.
CC SMART; SM00558; JmJc_1.
CC Zinc-finger; DNA-binding; Nuclear protein; Transcription regulation;
CC Metal-binding; Alternative splicing; Disease mutation.
CC ZN_FING 600 625
CC ZN_FING 1072 1126
CC VARSPLIC 1072 1126
CC Missing (in isoform Short).
CC /FTid=VSP_004276.
CC R -> Q (IN APL).
CC /FTid=VAR_005265.
CC T -> A (IN ALUNC).
CC /FTid=VAR_005266.
CC V -> D (IN ALUNC).
CC /FTid=VAR_005267.

```



```

FT DOMAIN 153 209 INTERMEMBRANE SPACE (POTENTIAL).
FT TRANSMEM 210 230 POTENTIAL.
FT DOMAIN 231 278 LUMENAL (POTENTIAL).
FT TRANSMEM 279 298 POTENTIAL.
FT DOMAIN 299 374 INTERMEMBRANE SPACE (POTENTIAL).
FT TRANSMEM 375 394 POTENTIAL.
FT DOMAIN 395 402 LUMENAL (POTENTIAL).
SQ SEQUENCE 402 AA; 43523 MW; 4970CCFF6577CB03 CRC64;

Query Match 84.8%; Score 39; DB 1; Length 402;
Best Local Similarity 87.5%; Pred No. 8.6; Mismatches 0; Indels 0; Gaps 0;
Matches 7; Conservative 1;

QY 2 EGEGRSGR 9
DB 88 EGEGRSGK 95

RESULT 2
COAT_MDV STANDARD; PRT; 260 AA.
AC P06946;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 01-JAN-1988 (Rel. 06, Last annotation update)
DE Coat protein.
OS Wheat dwarf virus (WDV).
OC Viruses; ssDNA viruses; Geminiviridae; Mastrevirus.
OC NCBI_TaxID=10834;
RN [1]
RP SEQUENCE FROM N.A.
RA McDowell S.W., McDonald H., Hamilton W.D.O., Coutts R.H.A.,
RT Buck K.W.;
RT "The nucleotide sequence of cloned wheat dwarf virus DNA.";
RL EMBO J. 4:2173-2180(1985).

-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; X02869; CAA26622.1; -.
PIR; A25005; VCCVWV.
DR InterPro; IPR000143; GemoCoat MSV.
DR InterPro; IPR000263; GV_A/BRI_coat.
DR Pfam; PF00844; Gemini_coat; 1.
DR PRINTS; PR00223; GEMCOATABRI.
DR ProDom; PD001964; GemoCoat_MSV; 1.
KW Coat protein.
SQ SEQUENCE 260 AA; 29408 MW; 45A2FFEA8416FCF7 CRC64;

Query Match 78.3%; Score 36; DB 1; Length 260;
Best Local Similarity 87.5%; Pred No. 19; Mismatches 0; Indels 0; Gaps 0;
Matches 7; Conservative 0;

QY 2 EGEGRSGR 9
DB 16 EGEGRSGK 23

RESULT 3
KDC2 DROME STANDARD; PRT; 502 AA.
AC P16912;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Protein kinase DC2 (EC 2.7.1.-).
GN PKA-C3 OR DC2.
OS Drosophila melanogaster (Fruit fly).

```

```

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
CX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Canton-S;
RX MEDLINE=89107990; PubMed=3215511;
RA Kalderson D., Rubin G.M.;
RT "Isolation and characterization of Drosophila cAMP-dependent protein
kinase genes.";
RL Genes Dev. 2:1539-1556(1988).
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- TISSUE SPECIFICITY: MORE ABUNDANT IN ADULT HEAD THAN ADULT BODY.
CC -1- DEVELOPMENTAL STAGE: IN EMBRYONS, PUPAE AND ADULTS.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CAMP SUBFAMILY.

-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; X16961; CAA34835.1; -.
HSP; P05132; IATP.
DR FlyBase; FBgn0000489; Pka-C3.
DR GO; GO:0008602; P:protein amino acid phosphorylation; IDA.
DR GO; GO:0006468; P:protein kinase.
DR InterPro; IPR000961; Pkinase C.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR002290; Ser Thr kinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00089; Pkinase; 1.
DR Pfam; PF00433; Pkinase C; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot kinase; 1.
DR SMART; SM00133; S_TK_X; 1.
DR SMART; SM00220; S_TK; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding.
FT DOMAIN 193 447 PROTEIN KINASE.
FT NP_BIND 199 207 ATP (BY SIMILARITY).
FT BINDING 222 222 ATP (BY SIMILARITY).
FT ACT_SITE 316 316 BY SIMILARITY.
SQ SEQUENCE 502 AA; 56960 MW; CAAC474A5A40193D CRC64;

Query Match 78.3%; Score 36; DB 1; Length 502;
Best Local Similarity 75.0%; Pred. No. 38; Mismatches 0; Indels 0; Gaps 0;
Matches 6; Conservative 2;

QY 2 EGEGRSGR 9
DB 136 DEGEGRNGR 143

RESULT 4
YLW2 CAEEL STANDARD; PRT; 286 AA.
AC P34396;
DT 01-FEB-1994 (Rel. 28, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein F10E9.2 in chromosome III.
GN F10E9.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OC NCBI_TaxID=6239;

```

A:Accession: S44801
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-304 <AND>
 A:Cross-references: EMBL:U10986; NID:gi56289; PID:gi56292
 C:Genetics:
 A:Introns: 13/2; 37/3; 99/1; 192/1; 233/3

Query Match 76.1%; Score 35; DB 2; Length 304;
 Best Local Similarity 81.8%; Pred. No. 69;
 Matches 9; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

Qy 1 SEEGEG--SGR 9
 |||||
 Db 220 SEEGEGERSGR 230

RESULT 13
 H86218
 protein F22013.12 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: H86218
 R:Neologs, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Dewar, A.R.; Creasy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.B.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; UID:21016719; PMID:11130712
 A:Accession: H86218
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-319 <STO>
 A:Cross-references: GB:AB005172; NID:99802577; PID:NAF99779.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: F22013.12
 A:Map position: 1
 C:Superfamily: Arabidopsis thaliana hypothetical protein F22013.12

Query Match 76.1%; Score 35; DB 2; Length 319;
 Best Local Similarity 77.8%; Pred. No. 72;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SEEGEGSGR 9
 :|||||
 Db 218 NEEGESGR 226

RESULT 14
 T00717
 hypothetical protein F22013.12 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 04-Mar-2000
 C:Accession: T00717
 R:Shinn, P.; Buehler, E.; Dewar, K.; Feng, J.; Kim, C.; Li, Y.; Sun, H.; Conway, A.; Con-
 cologis, A.; Ecker, J.R.
 submitted to the EMBL Data Library, April 1998
 A:Description: Genomic sequence for Arabidopsis thaliana BAC F22013.
 A:Reference number: Z14200
 A:Accession: T00717
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-364 <SHI>
 A:Cross-references: EMBL:AC003981; NID:G3063438; PID:G3063450; GSPDB:GN00059; ATSP:F2201
 C:Genetics:
 A:Gene: ATSP:F22013.12
 A:Map position: 1
 A:Introns: 44/2; 110/3; 157/3; 183/3; 202/2; 221/3; 263/3; 295/3; 319/3

C:Superfamily: Arabidopsis thaliana hypothetical protein F22013.12

Query Match 76.1%; Score 35; DB 2; Length 364;
 Best Local Similarity 77.8%; Pred. No. 82;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SEEGEGSGR 9
 :|||||
 Db 263 NEEGESGR 271

RESULT 15

T32155
 hypothetical protein C10F3.1 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T32155
 R:Davidson, S.; Wohldmann, P.
 submitted to the EMBL Data Library, September 1997
 A:Description: The sequence of C. elegans cosmid C10F3.
 A:Reference number: Z21127
 A:Accession: T32155
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-782 <DAV>
 A:Cross-references: EMBL:AF022968; PID:AA869883.1; GSPDB:GN00023; CBSP:C10F3.1
 C:Experimental source: strain Bristol N2; clone C10F3
 C:Genetics:
 A:Gene: CBSP:C10F3.1
 A:Map position: 5
 A:Introns: 52/3; 107/2; 224/1; 318/2; 378/2; 455/1; 657/2

Query Match 76.1%; Score 35; DB 2; Length 782;
 Best Local Similarity 87.5%; Pred. No. 1.8e+02;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SEEGEGSG 8
 :|||||
 Db 568 SERGESG 575

Search completed: December 11, 2003, 18:29:37
 Job time : 10.3333 secs

Db 221 EEEGEGGR 228
|||||
RESULT 3
T26747
hypothetical protein Y39A1B.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T26747
R:Wall, M.
submitted to the EMBL Data Library, January 1998
A:Reference number: Z20258
A:Accession: T26747
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1260 <WIL>
A:Cross-references: EMBL:AL021482; PIDN:CAA16340.2; GSPDB:GN000021; CESP:Y39A1B.3
A:Experimental source: clone Y39A1B
C:Genetics:
A:Gene: CESP:Y39A1B.3
A:Map position: 3
A:Introns: 92/1; 266/3; 330/2; 360/2; 413/3; 697/1; 917/2; 969/2; 1060/2; 1089/2; 1115/3
Query Match 80.4%; Score 37; DB 2; Length 1260;
Best Local Similarity 87.5%; Pred. No. 1.3e+02; Mismatches 1; Indels 0; Gaps 0;
Matches 7; Conservative 0;
QY 2 EEEGEGSGR 9
| | | | |
Db 82 EEEGEGSGR 89
| | | | |
RESULT 4
C89589
protein Y39A1B.3 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: C89589
R:Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A:Reference number: A75000; MUID:99069613; PMID:9851916
A>Note: see websites genome.wustl.edu/gsc/Celegans/ and www.sanger.ac.uk/Projects/C_ele
A>Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: C89589
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1276 <STO>
A:Cross-references: GB:chr_III; PIDN:CAA16340.1; PID:G3880800; GSPDB:GN000021; CESP:Y39A1
C:Genetics:
A:Gene: Y39A1B.3
A:Map position: 3
Query Match 80.4%; Score 37; DB 2; Length 1276;
Best Local Similarity 87.5%; Pred. No. 1.3e+02; Mismatches 1; Indels 0; Gaps 0;
Matches 7; Conservative 0;
QY 2 EEEGEGSGR 9
| | | | |
Db 82 EEEGEGSGR 89
| | | | |
RESULT 5
VCCVWV
coat protein - wheat dwarf virus
C:Species: wheat dwarf virus
C>Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 08-Apr-1994
C:Accession: A25005; A24356
R:MacDowell, S.W.; Macdonald, H.; Hamilton, W.D.O.; Coutts, R.H.A.; Buck, K.W.
EMBO J. 4, 2173-2180, 1985
A:Title: The nucleotide sequence of cloned wheat dwarf virus DNA.
A:Reference number: A91012

A:Accession: A25005
A:Molecule type: DNA
A:Residues: 1-260 <MAC>
A>Note: the authors translated the codon AAG for residue 5 as Arg
C:Superfamily: maize streak virus coat protein
C:Keywords: coat protein
Query Match 78.3%; Score 36; DB 1; Length 260;
Best Local Similarity 87.5%; Pred. No. 39;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 EEEGEGSGR 9
| | | | |
Db 16 EEEGEGSGR 23
| | | | |
RESULT 6
S49386
capsid protein V2 - wheat dwarf virus
C:Species: wheat dwarf virus
C>Date: 16-Feb-1995 #sequence_revision 12-May-1995 #text_change 20-Sep-1999
C:Accession: S49386
R:Bendahmane, M.; Schalk, H.J.; Gronenborn, B.
submitted to the EMBL Data Library, October 1994
A:Description: Identification and characterization of wheat dwarf virus (WDV) from Fra
A:Reference number: S49385
A:Accession: S49386
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-260 <BEN>
A:Cross-references: EMBL:X82104; NID:G557674; PIDN:CAA57624.1; PID:G557676
C:Superfamily: maize streak virus coat protein
Query Match 78.3%; Score 36; DB 2; Length 260;
Best Local Similarity 87.5%; Pred. No. 39;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 EEEGEGSGR 9
| | | | |
Db 16 EEEGEGSGR 23
| | | | |
RESULT 7
A10701
probable amino acid permease STY1743 [imported] - Salmonella enterica subsp. enterica
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A>Note: this species has also been called Salmonella typhi
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: A10701
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher
th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farra
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica ser
A:Reference number: AB0502; MUID:21534947; PMID:11677608
C:Accession: A10701
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-447 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD01986.1; PID:G16502825; GSPDB:GN00176
C:Genetics:
A:Gene: STY1743
Query Match 78.3%; Score 36; DB 2; Length 447;
Best Local Similarity 77.8%; Pred. No. 67;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 SEEEGEGSGR 9
| | | | |
Db 213 SEEEGEGSGR 221
| | | | |

```

; PRIOR APPLICATION NUMBER: 09/180997
; PRIOR FILING DATE: 1998-11-19
; PRIOR APPLICATION NUMBER: 09/267213
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 09/380137
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: 09/380138
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: 09/403297
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 09/423741
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: 09/709238
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 09/802706
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 09/872035
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 09/918565
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 09/924419
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: 09/927796
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/929404
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: 09/931836
; PRIOR FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: 09/941992
; PRIOR FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: 09/946374
; PRIOR FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: 10/001054
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 10/081056
; PRIOR FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: 10/119480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: PCT/US98/18824
; PRIOR FILING DATE: 1998-09-10

```

```

Query Match      78.3%; Score 36; DB 12; Length 839;
Best Local Similarity 75.0%; Pred. No. 9.1e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

OY      2 ECEGGGGR 9
      :|||||:
Db      817 DEEGGAGR 824

```

```

RESULT 15
US-10-211-884-36
; Sequence 36, Application US/10211884
; Publication No. US20030175900A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Marsters, Scot A.
; APPLICANT: Pan, James
; APPLICANT: Pitti, Robert M.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Smith, Victoria
; APPLICANT: Stone, Donna M.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF TUMOR
; FILE REFERENCE: P2931R1C1
; CURRENT APPLICATION NUMBER: US/10/211,884

```

```

; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: 60/014699
; PRIOR FILING DATE: 1996-04-01
; PRIOR APPLICATION NUMBER: 60/026943
; PRIOR FILING DATE: 1996-09-23
; PRIOR APPLICATION NUMBER: 60/059121
; PRIOR FILING DATE: 1997-07-17
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/062037
; PRIOR FILING DATE: 1997-10-10
; PRIOR APPLICATION NUMBER: 60/063755
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063045
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/066511
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066772
; PRIOR FILING DATE: 1997-11-24
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 258
; SEQ ID NO 36
; LENGTH: 839
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-211-884-36

```

```

Query Match      78.3%; Score 36; DB 12; Length 839;
Best Local Similarity 75.0%; Pred. No. 9.1e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

OY      2 ECEGGGGR 9
      :|||||:
Db      817 DEEGGAGR 824

```

```

Search completed: December 11, 2003, 18:38:03
Job time : 49.3333 secs

```

```

; LENGTH: 839
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-245-164-54

Query Match      78.3%; Score 36; DB 12; Length 839;
Best Local Similarity 75.0%; Pred. No. 9.1e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy      2 EGEGRSGR 9
Db      817 DEGEGR 824

```

```

RESULT 12
US-10-210-951-36
; Sequence 36, Application US/10210951
; Publication No. US2003017028A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Marsters, Scott A.
; APPLICANT: Pan, James
; APPLICANT: Pitti, Robert M.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Smith, Victoria
; APPLICANT: Stone, Donna M.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF TUMOR
; FILE REFERENCE: P2931R1C1
; CURRENT APPLICATION NUMBER: US/10/210,951
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: 60/014699
; PRIOR FILING DATE: 1996-04-01
; PRIOR APPLICATION NUMBER: 60/026943
; PRIOR FILING DATE: 1996-09-23
; PRIOR APPLICATION NUMBER: 60/059121
; PRIOR FILING DATE: 1997-07-17
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/062037
; PRIOR FILING DATE: 1997-10-10
; PRIOR APPLICATION NUMBER: 60/063755
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063045
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/066511
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066772
; PRIOR FILING DATE: 1997-11-24
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 258
; SEQ ID NO 36
; LENGTH: 839
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-210-951-36

Query Match      78.3%; Score 36; DB 12; Length 839;
Best Local Similarity 75.0%; Pred. No. 9.1e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy      2 EGEGRSGR 9
Db      817 DEGEGR 824

```

```

RESULT 13
US-10-244-972-54
; Sequence 54, Application US/10244972
; Publication No. US20030170803A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Baton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C75
; CURRENT APPLICATION NUMBER: US/10/244,972
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 54
; LENGTH: 839
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-244-972-54

Query Match      78.3%; Score 36; DB 12; Length 839;
Best Local Similarity 75.0%; Pred. No. 9.1e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy      2 EGEGRSGR 9
Db      817 DEGEGR 824

```

```

RESULT 14
US-10-197-942-54
; Sequence 54, Application US/10197942
; Publication No. US20030175862A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Baton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanabe, Colin

```

US-10-242-574-54

Query Match 78.3%; Score 36; DB 12; Length 839;
Best Local Similarity 75.0%; Pred. No. 9.1e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 EGECSGR 9
:||||:|
Db 817 DEGEAGR 824

RESULT 7

US-10-243-261-54
; Sequence 54, Application US/10243261
; Publication No. US20030138900A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C56

; CURRENT APPLICATION NUMBER: US/10/243,261

; CURRENT FILING DATE: 2002-09-13

; PRIOR APPLICATION NUMBER: 10/197942

; PRIOR FILING DATE: 2002-07-18

; PRIOR APPLICATION NUMBER: 60/059114

; PRIOR FILING DATE: 1997-09-17

; PRIOR APPLICATION NUMBER: 60/063046

; PRIOR FILING DATE: 1997-10-24

; PRIOR APPLICATION NUMBER: 60/065027

; PRIOR FILING DATE: 1997-11-10

; PRIOR APPLICATION NUMBER: 60/079689

; PRIOR FILING DATE: 1998-03-27

; PRIOR APPLICATION NUMBER: 60/086478

; PRIOR FILING DATE: 1998-05-22

; PRIOR APPLICATION NUMBER: 60/087607

; PRIOR FILING DATE: 1998-06-02

; PRIOR APPLICATION NUMBER: 60/089801

; PRIOR FILING DATE: 1998-06-18

; PRIOR APPLICATION NUMBER: 60/090557

; PRIOR FILING DATE: 1998-06-24

; PRIOR APPLICATION NUMBER: 60/090689

; PRIOR FILING DATE: 1998-06-25

; Remaining Prior Application data removed - See File Wrapper or PALM.
; SEQ ID NO 54

; LENGTH: 839

; TYPE: PRT

; ORGANISM: Homo Sapien

US-10-243-261-54

Query Match 78.3%; Score 36; DB 12; Length 839;
Best Local Similarity 75.0%; Pred. No. 9.1e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 EGECSGR 9
:||||:|
Db 817 DEGEAGR 824

RESULT 8

US-10-243-282-54
; Sequence 54, Application US/10243282

; Publication No. US20030138901A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C50

; CURRENT APPLICATION NUMBER: US/10/243,282

; CURRENT FILING DATE: 2002-09-13

; PRIOR APPLICATION NUMBER: 10/197942

; PRIOR FILING DATE: 2002-07-18

; PRIOR APPLICATION NUMBER: 60/059114

; PRIOR FILING DATE: 1997-09-17

; PRIOR APPLICATION NUMBER: 60/063046

; PRIOR FILING DATE: 1997-10-24

; PRIOR APPLICATION NUMBER: 60/065027

; PRIOR FILING DATE: 1997-11-10

; PRIOR APPLICATION NUMBER: 60/079689

; PRIOR FILING DATE: 1998-03-27

; PRIOR APPLICATION NUMBER: 60/086478

; PRIOR FILING DATE: 1998-05-22

; PRIOR APPLICATION NUMBER: 60/087607

; PRIOR FILING DATE: 1998-06-02

; PRIOR APPLICATION NUMBER: 60/089801

; PRIOR FILING DATE: 1998-06-18

; PRIOR APPLICATION NUMBER: 60/090557

; PRIOR FILING DATE: 1998-06-24

; PRIOR APPLICATION NUMBER: 60/090689

; PRIOR FILING DATE: 1998-06-25

; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116

; SEQ ID NO 54

; LENGTH: 839

; TYPE: PRT

; ORGANISM: Homo Sapien

US-10-243-282-54

Query Match 78.3%; Score 36; DB 12; Length 839;
Best Local Similarity 75.0%; Pred. No. 9.1e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 EGECSGR 9
:||||:|
Db 817 DEGEAGR 824

RESULT 9

US-10-243-402-54

; Sequence 54, Application US/10243402

; Publication No. US20030138902A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman

;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 39399
;; LENGTH: 85
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC002539.1
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.6
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.9
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.2
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.1
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.2
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.7
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.3
US-09-864-761-39399

Query Match 82.6%; Score 38; DB 9; Length 85;
Best Local Similarity 87.5%; Pred. No. 50;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 EGEGR 9
Db 27 EGEGR 34

RESULT 2

US-09-764-864-937
;; Sequence 937, Application US/09764864
;; Patent No. US20020132753A1
;; GENERAL INFORMATION:
;; APPLICANT: Rosen et al.
;; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;; FILE REFERENCE: PT223
;; CURRENT APPLICATION NUMBER: US/09/764,864
;; CURRENT FILING DATE: 2001-01-17
;; Prior application data removed - consult PALM or file wrapper
;; NUMBER OF SEQ ID NOS: 1792
;; SOFTWARE: Patent in Ver. 2.0
;; SEQ ID NO 937
;; LENGTH: 212
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-764-864-937

Query Match 82.6%; Score 38; DB 10; Length 212;
Best Local Similarity 87.5%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 EGEGR 9
Db 148 EGEGR 155

RESULT 3

US-10-237-496-54
;; Sequence 54, Application US/10237496
;; Publication No. US20030138896A1
;; GENERAL INFORMATION:
;; APPLICANT: Baker, Kevin

;; APPLICANT: Eaton, Dan
;; APPLICANT: Filvaroff, Ellen
;; APPLICANT: Goddard, Audrey
;; APPLICANT: Grimaldi, J. Christopher
;; APPLICANT: Gurney, Austin
;; APPLICANT: Smith, Victoria
;; APPLICANT: Stephan, Jean-Phillippe
;; APPLICANT: Watanabe, Colin
;; APPLICANT: Wood, William
;; APPLICANT: Zhang, Zemin
;; APPLICANT: Fong, Sherman
;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;; FILE REFERENCE: P3630R1C4
;; CURRENT APPLICATION NUMBER: US/10/237,496
;; CURRENT FILING DATE: 2002-09-06
;; PRIOR APPLICATION NUMBER: 10/197942
;; PRIOR FILING DATE: 2002-07-18
;; PRIOR APPLICATION NUMBER: 60/059114
;; PRIOR FILING DATE: 1997-09-17
;; PRIOR APPLICATION NUMBER: 60/063046
;; PRIOR FILING DATE: 1997-10-24
;; PRIOR APPLICATION NUMBER: 60/065027
;; PRIOR FILING DATE: 1997-11-10
;; PRIOR APPLICATION NUMBER: 60/079689
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/086478
;; PRIOR FILING DATE: 1998-05-22
;; PRIOR APPLICATION NUMBER: 60/087607
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/089801
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090689
;; PRIOR FILING DATE: 1998-06-25
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 116
;; SEQ ID NO 54
;; LENGTH: 839
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-237-496-54

Query Match 78.3%; Score 36; DB 12; Length 839;
Best Local Similarity 75.0%; Pred. No. 9.1e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 EGEGR 9
Db 817 DEGEGR 824

RESULT 4

US-10-242-074-54
;; Sequence 54, Application US/10242074
;; Publication No. US20030138897A1
;; GENERAL INFORMATION:
;; APPLICANT: Baker, Kevin
;; APPLICANT: Eaton, Dan
;; APPLICANT: Filvaroff, Ellen
;; APPLICANT: Goddard, Audrey
;; APPLICANT: Grimaldi, J. Christopher
;; APPLICANT: Gurney, Austin
;; APPLICANT: Smith, Victoria
;; APPLICANT: Stephan, Jean-Phillippe
;; APPLICANT: Watanabe, Colin
;; APPLICANT: Wood, William
;; APPLICANT: Zhang, Zemin
;; APPLICANT: Fong, Sherman
;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;; FILE REFERENCE: P3630R1C19

QY 1 SEEGEGSG 8
:|||||
Db 71 TEEGEGEG 78

Search completed: December 11, 2003, 18:30:37
Job time : 9.22222 secs

```

PCT-US91-02714-14
; Sequence 14, Application PC/TUS9102714
; GENERAL INFORMATION:
; APPLICANT: Wigler, Michael H.
; APPLICANT: Colicelli, John H.
; TITLE OF INVENTION: Cloning by Complementation and Related
; TITLE OF INVENTION: Processes
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: Two First National Plaza, 20 South Clark
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: PCT/US91/02714
; APPLICATION NUMBER: PCT/US91/02714
; FILING DATE: 19910419
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/511,715
; FILING DATE: 20-APR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Borun, Michael F.
; REGISTRATION NUMBER: 25447
; REFERENCE/DOCKET NUMBER: 27805/30197
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 405 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US91-02714-14

Query Match 76.1%; Score 35; DB 5; Length 405;
Best Local Similarity 75.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SEEGEGSG 8
Db 221 TQEGEGSG 228

RESULT 9
US-09-252-991A-18838
; Sequence 18838, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 14:
; LENGTH: 405
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30587
; Sequence 30587, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30587
; LENGTH: 694
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30587

Query Match 76.1%; Score 35; DB 4; Length 694;
Best Local Similarity 66.7%; Pred. No. 3.5e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SEEGEGSGR 9
Db 177 ADRGEGSGR 185

RESULT 11
US-09-287-354-2
; Sequence 2, Application US/09287354
; Patent No. 6348348
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Catherine C.
; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN
; FILE REFERENCE: Thompson-20263/0243435
; CURRENT APPLICATION NUMBER: US/09/287,354
; CURRENT FILING DATE: 1999-04-07
; EARLIER APPLICATION NUMBER: US 60/080,888
; EARLIER FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 984
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-287-354-2

Query Match 76.1%; Score 35; DB 4; Length 984;
Best Local Similarity 87.5%; Pred. No. 5e-02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SEEGEGSG 8
Db 333 SEEGEGSG 340
```

```
QY      1 SEEEGSG 8
      :|||||
Db      74 TQEGGSG 81

RESULT 2
US-09-522-955A-22
; Sequence 22, Application US/09522955A
; Patent No. 6500942
; GENERAL INFORMATION:
; APPLICANT: Tam, See-Ying
; APPLICANT: Tsai, Mindy
; APPLICANT: Galli, Stephen J.
; TITLE OF INVENTION: RIN2, A NOVEL INHIBITOR OF RAS-MEDICATED
; TITLE OF INVENTION: SIGNALING
; FILE REFERENCE: 1440.1089-004
; CURRENT APPLICATION NUMBER: US/09/522,955A
; CURRENT FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: PCI/US98/19056
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: US 08/942,819
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: US 60/058,520
; PRIOR FILING DATE: 1997-09-11
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-522-955A-22

Query Match      76.1%; Score 35; DB 4; Length 143;
Best Local Similarity 75.0%; Pred. No. 74;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 SEEEGSG 8
      :|||||
Db      74 TQEGGSG 81

RESULT 3
US-09-252-991A-17526
; Sequence 17526, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17526
; LENGTH: 167
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17526

Query Match      76.1%; Score 35; DB 4; Length 167;
Best Local Similarity 66.7%; Pred. No. 86;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 SEEEGSGR 9
      :|||||
Db      77 TQEGGSGR 85
```

```
RESULT 4
US-08-474-379C-14
; Sequence 14, Application US/08474379C
; Patent No. 5977305
; GENERAL INFORMATION:
; APPLICANT: Wigler, Michael H.
; APPLICANT: Colicelli, John J.
; TITLE OF INVENTION: CLONING BY COMPLEMENTATION AND RELATED
; TITLE OF INVENTION: PROCESSES
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive/6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,379C
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/511,715
; FILING DATE: 20-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/206,188
; FILING DATE: 01-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/688,352
; FILING DATE: 19-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 27866/32771
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 403 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-474-379C-14

Query Match      76.1%; Score 35; DB 2; Length 403;
Best Local Similarity 75.0%; Pred. No. 21e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 SEEEGSG 8
      :|||||
Db      219 TQEGGSG 226

RESULT 5
US-09-146-249A-14
; Sequence 14, Application US/09146249A
; Patent No. 6069240
; GENERAL INFORMATION:
; APPLICANT: Wigler, Michael H.
; APPLICANT: Colicelli, John J.
; TITLE OF INVENTION: Cloning by Complementation and Related
; TITLE OF INVENTION: Processes
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
```

OY 2 EREGSGR 9
|||
Db 148 EREGSGR 155

Search completed: December 11, 2003, 18:24:57
Job time : 27 secs

CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 85 AA;
 SQ Query Match 82.6%; Score 38; DB 23; Length 85;
 Best Local Similarity 87.5%; Pred. No. 28;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 EGEQSGSR 9
 |||||
 Db 27 EGEQSGSR 34

RESULT 14
 ABG26522
 ID ABG26522 standard; Protein; 152 AA.

XX AC ABG26522;
 XX DT 18-FEB-2002 (first entry)
 XX DE Novel human diagnostic protein #26513.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX PI Dmanac RT, Liu C, Tang YT;

XX PR 2001-639362/73.

XX DR N-PSDB; AAS90709.

XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -

XX Claim 20; SEQ ID No 56881; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG0377 represent novel human
 CC diagnostic amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 152 AA;
 SQ Query Match 82.6%; Score 38; DB 22; Length 152;
 Best Local Similarity 87.5%; Pred. No. 51;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 EGEQSGSR 9
 |||||
 Db 75 EGEQSGSR 82

RESULT 15

AAU15984

ID AAU15984 standard; Protein; 212 AA.

XX AC AAU15984;

XX DT 07-NOV-2001 (first entry)

XX DE Human novel secreted protein, Seq ID 937.

XX KW Human; immunosuppressive; antiarthritic; antirheumatic;
 KW cytosolic; cardiant; vasotropic; cerebroprotective; neurotropic;
 KW neuroprotective; antibacterial; virucide; fungicide; ophthalmological;
 KW vulnary; secreted protein; rheumatoid arthritis;
 KW hyperproliferative disorder; cardiovascular disorder; cardiac arrest;
 KW cerebrovascular disorder; cerebral ischaemia; angiogenesis;
 KW nervous system disorder; Alzheimer's disease; infection; ocular disorder;
 KW corneal infection; wound healing; epithelial cell proliferation;
 KW skin ageing; food additive; preservative; antiproliferative.

XX OS Homo sapiens.

XX PN WO200155322-A2.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US01341.

XX PR 31-JAN-2000; 2000US-0179065.

XX PR 04-FEB-2000; 2000US-0180628.

XX PR 24-FEB-2000; 2000US-0184664.

XX PR 02-MAR-2000; 2000US-0186350.

XX PR 16-MAR-2000; 2000US-0189874.

XX PR 17-MAR-2000; 2000US-0190076.

XX PR 18-APR-2000; 2000US-0198123.

XX PR 19-MAY-2000; 2000US-0205515.

XX PR 07-JUN-2000; 2000US-0209467.

XX PR 28-JUN-2000; 2000US-0214886.

XX PR 30-JUN-2000; 2000US-0215135.

XX PR 07-JUL-2000; 2000US-0216647.

XX PR 07-JUL-2000; 2000US-0216880.

XX PR 11-JUL-2000; 2000US-0217487.

XX PR 11-JUL-2000; 2000US-0217496.

XX PR 14-JUL-2000; 2000US-0218230.

XX PR 26-JUL-2000; 2000US-0220963.

XX PR 26-JUL-2000; 2000US-0220964.

XX PR 14-AUG-2000; 2000US-0224518.

XX PR 14-AUG-2000; 2000US-0224519.

XX PR 14-AUG-2000; 2000US-0225213.

XX PR 14-AUG-2000; 2000US-0225214.

XX PR 14-AUG-2000; 2000US-0225266.

XX PR 14-AUG-2000; 2000US-0225267.

XX PR 14-AUG-2000; 2000US-0225268.

XX PR 14-AUG-2000; 2000US-0225270.

XX PR 14-AUG-2000; 2000US-0225447.

XX PR 14-AUG-2000; 2000US-0225757.

XX PR 14-AUG-2000; 2000US-0225758.

XX PR 14-AUG-2000; 2000US-0225759.

XX PR 18-AUG-2000; 2000US-0226279.

XX PR 22-AUG-2000; 2000US-0226681.

XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-483446/52.
 XX Single exon nucleic acid probes for analyzing gene expression in human
 XX brains -
 XX
 XX Example 4; SEQ ID NO: 32041; 650pp + Sequence Listing; English.
 XX
 XX The present invention provides a number of single exon nucleic acid
 XX probes which are derived from genomic sequences expressed in the human
 XX brain. They can be used to measure gene expression in brain cell samples,
 XX which may enable the diagnosis and improved treatment of nervous system
 XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 XX epilepsy and cancers. The present sequence is a protein encoded by one of
 XX the probes of the invention.
 XX
 XX SQ Sequence 85 AA;
 XX
 XX Query Match 82.6%; Score 38; DB 22; Length 85;
 XX Best Local Similarity 87.5%; Pred. No. 28;
 XX Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 XX
 XX QY 2 BEGEGSGR 9
 XX |||||
 XX 27 BEGEGKGR 34
 XX
 XX
 XX RESULT 10
 XX AAM72531
 XX ID AAM72531 standard; Protein; 85 AA.
 XX AC AAM72531;
 XX
 XX DT 06-NOV-2001 (first entry)
 XX
 XX DE Human bone marrow expressed probe encoded protein SEQ ID NO: 32837.
 XX
 XX KW Human; bone marrow expressed exon; gene expression analysis; probe;
 XX microarray; cancer; leukaemia; lymphoma; myeloma.
 XX
 XX OS Homo sapiens.
 XX
 XX PN WO200157276-A2.
 XX
 XX PD 09-AUG-2001.
 XX
 XX PF 30-JAN-2001; 2001WO-US00668.
 XX
 XX PR 04-FEB-2000; 2000US-0180312.
 XX PR 26-MAY-2000; 2000US-0207456.
 XX PR 30-JUN-2000; 2000US-0608408.
 XX PR 03-AUG-2000; 2000US-0632366.
 XX PR 21-SEP-2000; 2000US-0234687.
 XX PR 27-SEP-2000; 2000US-0236359.
 XX PR 04-OCT-2000; 2000GB-0024263.
 XX
 XX PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 XX PD 2001-488900/53.
 XX
 XX PF Human genome-derived single exon nucleic acid probes useful for
 XX analyzing gene expression in human bone marrow -
 XX
 XX PS Example 4; SEQ ID NO: 32837; 650pp + Sequence Listing; English.
 XX
 XX The present invention provides a number of single exon nucleic acid
 XX probes which are derived from genomic sequences expressed in the human
 XX bone marrow. They can be used to measure gene expression in bone marrow
 XX samples, which may enable the improved diagnosis and treatment of cancers

CC such as lymphoma, leukaemia and myeloma. The present sequence is a
 CC protein encoded by one of the probes of the invention.
 XX
 XX SQ Sequence 85 AA;
 XX
 XX Query Match 82.6%; Score 38; DB 22; Length 85;
 XX Best Local Similarity 87.5%; Pred. No. 28;
 XX Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 XX
 XX QY 2 BEGEGSGR 9
 XX |||||
 XX 27 BEGEGKGR 34
 XX
 XX
 XX RESULT 11
 XX AAM19618
 XX ID AAM19618 standard; Protein; 85 AA.
 XX AC AAM19618;
 XX
 XX DT 12-OCT-2001 (first entry)
 XX
 XX DE Peptide #6052 encoded by probe for measuring cervical gene expression.
 XX
 XX KW Probe; human; microarray; gene expression; cervical epithelial cell;
 XX cervical cancer.
 XX
 XX OS Homo sapiens.
 XX
 XX PN WO200157278-A2.
 XX
 XX PD 09-AUG-2001.
 XX
 XX PF 30-JAN-2001; 2001WO-US00670.
 XX
 XX PR 04-FEB-2000; 2000US-0180312.
 XX PR 26-MAY-2000; 2000US-0207456.
 XX PR 30-JUN-2000; 2000US-0608408.
 XX PR 03-AUG-2000; 2000US-0632366.
 XX PR 21-SEP-2000; 2000US-0234687.
 XX PR 27-SEP-2000; 2000US-0236359.
 XX PR 04-OCT-2000; 2000GB-0024263.
 XX
 XX PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 XX PD 2001-488901/53.
 XX
 XX PF Human genome-derived single exon nucleic acid probes useful for
 XX analyzing gene expression in human cervical epithelial cells -
 XX
 XX PS Claim 27; SEQ ID NO 24444; 487pp; English.
 XX
 XX The present invention relates to human single exon nucleic acid probes
 XX (SENPs; see AAI10068-AA128459). The present sequence is a peptide encoded
 XX by one such probe. The SENPs are derived from human Hela cells. The SENPs
 XX can be used to produce a single exon microarray, which can be used for
 XX measuring human gene expression in a sample derived from human cervical
 XX epithelial cells. By measuring gene expression, the probes are therefore
 XX useful in grading and/or staging of diseases of the cervix, notably
 XX cervical cancer.
 XX
 XX Note: The sequence data for this patent did not form part of the printed
 XX specification, but was obtained in electronic format directly from WIPO
 XX at ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX SQ Sequence 85 AA;
 XX
 XX Query Match 82.6%; Score 38; DB 22; Length 85;
 XX Best Local Similarity 87.5%; Pred. No. 28;
 XX Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 XX
 XX QY 2 BEGEGSGR 9

OS Homo sapiens.
 PN WO200150366-A2.
 XX 29-NOV-2001.
 PD
 XX 24-MAY-2001; 2001WO-US17076.
 PF
 XX 24-MAY-2000; 2000US-206690P.
 PR
 XX (CURA-) CURAGEN CORP.
 PA
 XX Leach MD, Shimkets RA;
 PI
 XX WPI; 2002-106200/14.
 DR
 DR N-PSDB; ABN76558.
 XX
 XX Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and disorders related to organ
 PT transplantation
 XX
 PS Claim 10; Page 1011-1012; 2508pp; English.
 XX
 CC Sequences ABP31028-ABP35561 represent 4534 novel human proteins
 CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-
 CC ABN75587 represent cDNAs encoding them. The invention also encompasses
 CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
 CC referred to as ORFX) proteins, polynucleotides at least 85% identical to
 CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
 CC polynucleotides, the recombinant production of ORFX proteins, antibodies
 CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and
 CC polypeptides, methods of screening for modulators of ORFX expression or
 CC activity, and methods of screening individuals for a predisposition to an
 CC ORFX-associated disorder. The ORFX proteins of the invention have a wide
 CC range of biological activities, such as cytokine, cell proliferation,
 CC cell differentiation, immune modulation, haematopoiesis regulation,
 CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
 CC chemokinetic activity, haemostatic activity, thrombolytic activity,
 CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,
 CC and antifertility activity, and may also be involved in the determination
 CC of bodily characteristics, fertility and behaviour. ORFX proteins,
 CC nucleic acids and antibodies may be used in the treatment of cancers,
 CC other proliferative disorders such as psoriasis and benign tumours,
 CC neurological disorders such as epilepsy and Alzheimer's disease,
 CC cardiovascular diseases, immune system disorders, disorders related to
 CC organ transplantation, disorders of tissue growth and regeneration,
 CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
 CC storage disease, and infectious diseases caused by viral, bacterial,
 CC fungal and other pathogens. ORFX nucleic acids may also be used as a
 CC source of primers and probes, in the detection of ORFX genomic sequences
 CC or transcripts, in the identification and cloning of homologous
 CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
 CC nucleic acids may additionally be used to produce transgenic animals
 CC which may be useful for studying the function and/or activity of ORFX
 CC protein, and in drug screening. The ORFX proteins may also be used as
 CC immunogens to generate specific antibodies, which are useful in the
 CC diagnosis, treatment and monitoring of ORFX-associated diseases.
 XX
 SQ Sequence 75 AA;
 Query Match 82.6%; Score 38; DB 23; Length 75;
 Best Local Similarity 87.5%; Pred. No. 25;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 EGEQSGSR 9
 |||||
 DB 22 EGEQGGGR 29
 RESULT 6
 ABG54222
 ID ABG54222 standard; Peptide; 85 AA.

XX ABG54222;
 AC 25-FEB-2003 (first entry)
 DT
 XX Human liver peptide, SEQ ID No 32870.
 DE
 XX Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
 KW hypercholesterolaemia; coronary heart disease.
 XX
 OS Homo sapiens.
 XX WO200157273-A2.
 PN
 XX 09-AUG-2001.
 PD
 XX 30-JAN-2001; 2001WO-US00664.
 PF
 XX 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0609408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 PI
 XX WPI; 2001-488998/53.
 DR
 XX Human genome-derived single exon nucleic acid probes useful for
 PT analysing gene expression in human adult liver -
 PT
 XX Claim 27; SEQ ID No 32870; 658pp; English.
 XX
 CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
 CC measuring human gene expression in a sample derived from human adult
 CC liver, comprising one of 13109 defined nucleotide sequences given in the
 CC specification (or complements/ fragments). The probe hybridises at high
 CC stringency to a nucleic acid molecule expressed in the human adult
 CC liver. (I) may be used for predicting, measuring and displaying gene
 CC expression in samples derived from human adult liver. The genes
 CC identified may be involved in genetic liver diseases such as cirrhosis,
 CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which
 CC is associated with coronary heart disease. ABG47348-ABG59930 represent
 CC human liver single exon encoded peptides of the invention.
 CC Note: The sequence information for this patent does not appear in the
 CC printed specification but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 85 AA;
 Query Match 82.6%; Score 38; DB 22; Length 85;
 Best Local Similarity 87.5%; Pred. No. 25;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 EGEQSGSR 9
 |||||
 DB 27 EGEQGGGR 34
 RESULT 7
 ABG39272
 ID ABG39272 standard; Peptide; 85 AA.
 XX
 XX ABG39272;
 AC
 XX 04-FEB-2002 (first entry)
 DT
 XX Peptide #6778 encoded by human foetal liver single exon probe.
 DE
 XX

XX WO9958098-A2.
 XX 18-NOV-1999.
 XX 12-MAY-1999; 99WO-US10549.
 XX 12-MAY-1998; 98US-0085197.
 XX 05-MAY-1999; 99US-0085197.
 XX (BOCK/) BOCK S C.
 XX (PICA/) PICARD V.
 XX (ZEND/) ZENDEHROUH P.
 XX Bock SC, Picard V, Zendeherouh P;
 XX WPI; 2000-116274/10.
 XX New modified human antithrombin III compounds, used for treating e.g.
 XX sepsis, trauma, acute respiratory distress syndrome, restenosis,
 XX thrombosis, thromboembolism or stroke -
 XX Claim 13; Page 57; 75pp; English.
 XX The present sequence is from an antithrombin III (ATIII) variant, Bb.D
 XX derived from human ATIII.N135A cDNA insert of the pBlueBac baculovirus
 XX expression construct and comprises residues 385-393. The variant has
 XX improved resistance to elastase and IGG-activated neutrophils while
 XX retains anti-thrombin and anti-factor Xa activities. It may be
 XX expressed as glycoforms with enhanced heparin affinity which target the
 XX blood vessel wall more efficiently than ATIIIs with normal heparin
 XX affinity. The modified ATIIIs can be used to treat thrombin activation-
 XX related pathological symptoms due to sepsis, trauma, acute
 XX respiratory distress syndrome, restenosis, thrombosis, thromboembolism
 XX and stroke. It can also be used to reduce the risk of reocclusion
 XX and restenosis in percutaneous transluminal coronary angioplasty,
 XX thrombosis associated with surgery, ischaemia/reperfusion injury, and
 XX coagulation abnormalities in cancer or surgical patients.
 XX Sequence 9 AA;
 XX Query Match 100.0%; Score 46; DB 21; Length 9;
 XX Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SBECEGSGR 9
 DB 1 SBECEGSGR 9
 RESULT 2
 AAY4470
 ID AAY44470 standard; peptide; 9 AA.
 XX AC AAY44470;
 XX 27-MAR-2000 (first entry)
 XX Human antithrombin III variant Bb.B (385-393).
 XX Human; antithrombin III; ATIII variant Bb.B; elastase-resistant;
 XX IGG activated neutrophil resistant; anti-thrombin activity; heparin;
 XX anti-factor Xa activity; blood clotting disorder; sepsis; trauma; stroke;
 XX thrombin activation-related pathological symptom; restenosis; thrombosis;
 XX acute respiratory distress syndrome; thromboembolism; reocclusion.
 XX Homo sapiens.
 OS Synthetic.
 XX Key Location/Qualifiers
 XX Misc-difference 3
 FT Misc-difference 4 /note= "ATIII.N135A Ala at 387 is substituted by Glu"
 FT

FT Misc-difference 5 /note= "ATIII.N135A Val at 388 is substituted by Gly"
 FT Misc-difference 6 /note= "ATIII.N135A Val at 389 is substituted by Glu"
 FT Misc-difference 7 /note= "ATIII.N135A Ile at 390 is substituted by Gly"
 FT Misc-difference 8 /note= "ATIII.N135A Ala at 391 is substituted by Ser"
 XX WO9958098-A2.
 XX 18-NOV-1999.
 XX 12-MAY-1999; 99WO-US10549.
 XX 12-MAY-1998; 98US-0085197.
 XX 05-MAY-1999; 99US-0085197.
 XX (BOCK/) BOCK S C.
 XX (PICA/) PICARD V.
 XX (ZEND/) ZENDEHROUH P.
 XX Bock SC, Picard V, Zendeherouh P;
 XX WPI; 2000-116274/10.
 XX New modified human antithrombin III compounds, used for treating e.g.
 XX sepsis, trauma, acute respiratory distress syndrome, restenosis,
 XX thrombosis, thromboembolism or stroke -
 XX Claim 13; Page 57; 75pp; English.
 XX The present sequence is from an antithrombin III (ATIII) variant, Bb.B
 XX derived from human ATIII.N135A cDNA insert of the pBlueBac baculovirus
 XX expression construct and comprises residues 385-393. The variant has
 XX improved resistance to elastase and IGG-activated neutrophils while
 XX retains anti-thrombin and anti-factor Xa activities. It may be
 XX expressed as glycoforms with enhanced heparin affinity which target the
 XX blood vessel wall more efficiently than ATIIIs with normal heparin
 XX affinity. The modified ATIIIs can be used to treat thrombin activation-
 XX related pathological symptoms due to sepsis, trauma, acute
 XX respiratory distress syndrome, restenosis, thrombosis, thromboembolism
 XX and stroke. It can also be used to reduce the risk of reocclusion
 XX and restenosis in percutaneous transluminal coronary angioplasty,
 XX thrombosis associated with surgery, ischaemia/reperfusion injury, and
 XX coagulation abnormalities in cancer or surgical patients.
 XX Sequence 9 AA;
 XX Query Match 87.0%; Score 40; DB 21; Length 9;
 XX Best Local Similarity 88.9%; Pred. No. 9.3e+05;
 XX Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 SBECEGSGR 9
 DB 1 SBECEGSGR 9
 RESULT 3
 AAY4471
 ID AAY44471 standard; peptide; 9 AA.
 XX AC AAY44471;
 XX 27-MAR-2000 (first entry)
 XX Human antithrombin III variant Bb.C (residues 385-393).
 XX Human; antithrombin III; ATIII variant Bb.C; elastase-resistant;
 XX IGG activated neutrophil resistant; anti-thrombin activity; heparin;
 XX anti-factor Xa activity; blood clotting disorder; sepsis; trauma; stroke;
 XX thrombin activation-related pathological symptom; restenosis; thrombosis;
 XX acute respiratory distress syndrome; thromboembolism; reocclusion.
 XX


```
DR SMART; SMO0398; HMG; 1.
FT NON TER 329 329
SQ SEQUENCE 329 AA; 36938 MW; 954C2A601BEA11D7 CRC64;

Query Match 84.8%; Score 39; DB 11; Length 329;
Best Local Similarity 77.8%; Pred. No. 26;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 STEGEGSGR 9
Db 11 STEGEGDGK 19

RESULT 9
Q8C7E0 PRELIMINARY; PRT; 450 AA.
AC Q8C7E0;
DT 01-MAR-2003 (TREMELrel. 23, Created)
DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE Bobby sox homolog (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Pancreas;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK050488; BAC34285.1; -.
FT NON TER 450 450
SQ SEQUENCE 450 AA; 50804 MW; 9AB464903212A01C CRC64;

Query Match 84.8%; Score 39; DB 11; Length 450;
Best Local Similarity 77.8%; Pred. No. 37;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 STEGEGSGR 9
Db 11 STEGEGDGK 19

RESULT 10
Q8V150 PRELIMINARY; PRT; 453 AA.
AC Q8V150;
DT 01-MAR-2002 (TREMELrel. 20, Created)
DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMELrel. 21, Last annotation update)
DE HMG-box containing protein (Fragment).
GN HBP2
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA Sanchez-Diaz A., Blanco M., Jones N., Moreno S.;
RA "HBP2: A new mammalian protein that regulates G1/S transition in
RT fission yeast.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF276950; AAL68986.1; -.
DR InterPro; IPR000910; HMG_12_box.
DR Pfam; PF00505; HMG_box; I.
DR SMART; SMO0398; HMG; 1.
FT NON TER 453 453
SQ SEQUENCE 453 AA; 51160 MW; 7BBD239DD4649032 CRC64;

Query Match 84.8%; Score 39; DB 11; Length 453;
Best Local Similarity 77.8%; Pred. No. 37;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 STEGEGSGR 9
Db 11 STEGEGDGK 19

RESULT 11
Q8V149 PRELIMINARY; PRT; 455 AA.
AC Q8V149;
DT 01-MAR-2002 (TREMELrel. 20, Created)
DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMELrel. 21, Last annotation update)
DE HMG-box containing protein (Fragment).
GN HBP2
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Mammary gland;
RA Sanchez-Diaz A., Blanco M., Jones N., Moreno S.;
RA "HBP2: A new mammalian protein that regulates G1/S transition in
RT fission yeast.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF276951; AAL68987.1; -.
DR InterPro; IPR000910; HMG_12_box.
DR Pfam; PF00505; HMG_box; I.
DR SMART; SMO0398; HMG; 1.
FT NON TER 455 455
SQ SEQUENCE 455 AA; 51417 MW; A01D4BBD239DD464 CRC64;

Query Match 84.8%; Score 39; DB 11; Length 455;
Best Local Similarity 77.8%; Pred. No. 37;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 STEGEGSGR 9
Db 11 STEGEGDGK 19

RESULT 12
Q8CDV1 PRELIMINARY; PRT; 712 AA.
AC Q8CDV1;
DT 01-MAR-2003 (TREMELrel. 23, Created)
DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE Bobby sox homolog (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK029532; BAC26500.1; -.
FT NON TER 712 712
SQ SEQUENCE 712 AA; 79620 MW; DA1DEBB0C82B2A14 CRC64;

Query Match 84.8%; Score 39; DB 11; Length 712;
Best Local Similarity 77.8%; Pred. No. 61;
```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=98116662; PubMed=945484;
 RA Seki N., Ohira M., Nagase T., Ishikawa K., Miyajima N., Nakajima D.,
 RA Nomura N., Ohara O.;
 RT "Characterization of cDNA clones in size-fractionated cDNA libraries
 RT from human brain.";
 RL DNA Res. 4:345-349 (1997).
 DR ENBL; AB007924; BAA32300.2; --
 DR InterPro; IPR000328; PA_PTPase.
 DR Pfam; PF01569; PAP2; 1.
 DR SMART; SM00014; acidPPc; 1.
 KW Hypothetical protein.
 FT NON TER 1
 SQ SEQUENCE 607 AA; 66416 MW; C06BEF10106E625D CRC64;

 Query Match 89.1%; Score 41; DB 4; Length 607;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 STEGESG 8
 DB 482 STEGESG 489

 RESULT 3
 ID Q8N3F1 PRELIMINARY; PRT; 653 AA.
 AC Q8N3F1;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein (Fragment).
 DE Hypothetical protein (Fragment).
 GN DKFZP761A0623
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Amalgama;
 RA Bloeker H., Boecher M., Brandt P., Mewes H.W., Weill B., Wiemann S.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR ENBL; AL834390; CA039052.1; --
 DR InterPro; IPR000326; PA_PTPase.
 DR Pfam; PF01569; PAP2; 1.
 DR SMART; SM00014; acidPPc; 1.
 KW Hypothetical protein.
 FT NON TER 1
 SQ SEQUENCE 653 AA; 71417 MW; 58F7EAB9A96BA67F CRC64;

 Query Match 89.1%; Score 41; DB 4; Length 653;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 STEGESG 8
 DB 528 STEGESG 535

 RESULT 4
 ID Q8BV73 PRELIMINARY; PRT; 717 AA.
 AC Q8BV73;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical acid phosphatase/vanadium-dependent haloperoxidase
 DE structure containing protein.

OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Spinal cord;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573 (2002).
 DR ENBL; AK079635; BAC37711.1; --
 KW Hypothetical protein.
 SQ SEQUENCE 717 AA; 78502 MW; 430D63F20FB0C778 CRC64;

 Query Match 89.1%; Score 41; DB 11; Length 717;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 STEGESG 8
 DB 592 STEGESG 599

 RESULT 5
 ID Q8BXK2 PRELIMINARY; PRT; 766 AA.
 AC Q8BXK2;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical acid phosphatase/vanadium-dependent haloperoxidase
 DE structure containing protein.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Medulla oblongata;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573 (2002).
 DR ENBL; AK046782; BAC32865.1; --
 KW Hypothetical protein.
 SQ SEQUENCE 766 AA; 83304 MW; 404E79D22E2D29C0 CRC64;

 Query Match 89.1%; Score 41; DB 11; Length 766;
 Best Local Similarity 100.0%; Pred. No. 28;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 STEGESG 8
 DB 641 STEGESG 648

 RESULT 6
 ID Q9SMX8 PRELIMINARY; PRT; 167 AA.
 AC Q9SMX8;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Squamosa promoter binding protein-like 4 (Fragment).
 GN SPL4.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

```

CC -|- CATALYTIC ACTIVITY: 10-formyltetrahydrofolate + L-methionyl-
CC tRNA(Met) + H(2)O = tetrahydrofolate + N-formylmethionyl-
CC tRNA(Met).
CC
CC -|- SIMILARITY: Belongs to the fnt family.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE009665; AAL53506.1; -
CC EMBL; AE014596; AAN34201.1; -
CC DR EMBL; AG3542; AG3542.
CC DR TIGR; BRA1034; -
CC DR HAMAP; MF_00182; -; 1.
CC DR InterPro; IPR005794; Fmt.
CC DR InterPro; IPR005793; Formyl_trans_C.
CC DR InterPro; IPR002376; formyl_transf.
CC DR InterPro; IPR001555; GART.
CC DR Pfam; PF02911; formyl_trans_C; 1.
CC DR Pfam; PF00551; formyl_transf; 1.
CC DR TIGRfam; TIGR00460; Fmt; 1.
CC DR PROSITE; PS00373; GART; FALSE NEG.
CC KW Transferase; Methyltransferase; Protein biosynthesis;
CC Complete proteome.
CC FT BINDING 110 113 TETRAHYDROFOLATE (THF) (BY SIMILARITY).
CC SEQUENCE 306 AA; 32772 MW; E659DF13AEB6P2F7 CRC64;
CC
CC Query Match 71.7%; Score 33; DB 1; Length 306;
CC Best Local Similarity 87.5%; Pred. No. 61;
CC Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
CC
CC QY 1 STEGEGSG 8
CC |||||
CC 252 STLGECSG 259
CC
CC RESULT 14
CC SYNT2_MOUSE STANDARD; PRT; 422 AA.
CC AC P46037;
CC DT 01-NOV-1995 (Rel. 32, Created)
CC DT 01-NOV-1995 (Rel. 32, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE Synaptotagmin II (SyntII).
CC GN SYNT2.
CC OS Mus musculus (Mouse).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CC OX NCBI_TaxID=10090;
CC RN [1]
CC SEQUENCE FROM N.A.
CC RP MEDLINE=95050743; PubMed=7961887;
CC RX Fukuda M., Aruga J., Niinobe M., Aimoto S., Mikoshiba K.;
CC RT "Inositol-1,3,4,5-tetrakisphosphate binding to C2B domain of
CC IP4BP/synaptotagmin II.";
CC RL J. Biol. Chem. 269:29206-29211(1994).
CC RN [2]
CC SEQUENCE FROM N.A.
CC RP Adachi R., Teich A.H., Nigam R.;
CC RA "Genomic structure of the murine Synt2 gene.";
CC RT Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -|- FUNCTION: MAY HAVE A REGULATORY ROLE IN THE MEMBRANE INTERACTIONS
CC DURING TRAFFICKING OF SYNAPTIC VESICLES AT THE ACTIVE ZONE OF THE
CC SYNAPSE. IT BINDS ACIDIC PHOSPHOLIPIDS WITH A SPECIFICITY THAT
CC REQUIRES THE PRESENCE OF BOTH AN ACIDIC HEAD GROUP AND A DIACYL
CC BACKBONE.
CC -|- SUBUNIT: Homotetramer (Probable). Interacts with stonin 2.
CC -|- SUBCELLULAR LOCATION: SYNAPTIC VESICLES AND CHROMAFFIN GRANULES.
CC -|- DOMAIN: The first C2 domain mediates Ca(2+)-dependent phospholipid

```

```

CC binding.
CC -|- DOMAIN: The second C2 domain mediates interaction with Stonin 2.
CC -|- SIMILARITY: Contains 2 C2 domains.
CC -|- SIMILARITY: BELONGS TO THE SYNAPTOTAGMIN FAMILY.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D37793; BAA07041.1; -
CC EMBL; AF257303; AAF68987.1; -
CC DR EMBL; AF257304; AAF68988.1; -
CC DR HSSP; P21707; 1BYN.
CC DR WGD; MGI:99666; Syt2.
CC DR InterPro; IPR000008; C2.
CC DR InterPro; IPR002149; LRI.
CC DR InterPro; IPR001565; Synaptotagmin.
CC DR Pfam; PF00168; C2; 2.
CC DR PRINTS; PR00360; C2DOMAIN.
CC DR PRINTS; PR00399; SYNAPTOTAGMN.
CC DR SMART; SM00239; C2; 2.
CC DR PROSITE; PS00499; C2_DOMAIN_1; 2.
CC DR PROSITE; PS00004; C2_DOMAIN_2; 2.
CC KW Transmembrane; Repeat; Synapse; Glycoprotein.
CC FT DOMAIN 1 60 VESICULAR (POTENTIAL).
CC FT TRANSMEM 61 87 POTENTIAL.
CC FT DOMAIN 88 422 CYTOPLASMIC (POTENTIAL).
CC FT DOMAIN 136 382 PHOSPHOLIPID BINDING (PROBABLE).
CC FT DOMAIN 156 245 C2 DOMAIN 1.
CC FT DOMAIN 287 378 C2 DOMAIN 2.
CC FT CARBOHYD 32 32 N-LINKED (GLCNAC... ) (POTENTIAL).
CC SQ SEQUENCE 422 AA; 47262 MW; B4BD13FF70E0481B CRC64;
CC
CC Query Match 71.7%; Score 33; DB 1; Length 422;
CC Best Local Similarity 85.7%; Pred. No. 86;
CC Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
CC
CC QY 2 TEGEGSG 8
CC |||||
CC 128 TEGEGG 134
CC
CC RESULT 15
CC SYNT2_RAT STANDARD; PRT; 422 AA.
CC AC P29101;
CC DT 01-DEC-1992 (Rel. 24, Created)
CC DT 01-DEC-1992 (Rel. 24, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE Synaptotagmin II (SyntII).
CC GN SYNT2.
CC OS Rattus norvegicus (Rat).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CC OX NCBI_TaxID=10116;
CC RN [1]
CC SEQUENCE FROM N.A.
CC RP MEDLINE=91310620; PubMed=1856191;
CC RA Geppert M., Archer B.T. III, Suedhof T.C.;
CC RT "Synaptotagmin II: A novel differentially distributed form of
CC synaptotagmin.";
CC RL J. Biol. Chem. 266:13548-13552(1991).
CC -|- FUNCTION: MAY HAVE A REGULATORY ROLE IN THE MEMBRANE INTERACTIONS
CC DURING TRAFFICKING OF SYNAPTIC VESICLES AT THE ACTIVE ZONE OF THE
CC SYNAPSE. IT BINDS ACIDIC PHOSPHOLIPIDS WITH A SPECIFICITY THAT
CC REQUIRES THE PRESENCE OF BOTH AN ACIDIC HEAD GROUP AND A DIACYL
CC BACKBONE.
CC -|- SUBUNIT: Homotetramer (Probable). Interacts with Stonin 2 (By
CC similarity).

```

01-OCT-1993 (Rel. 27, Last sequence update)
 15-SEP-2003 (Rel. 42, Last annotation update)
 Ca(2+)/calmodulin-responsive adenylylate cyclase (EC 4.6.1.1) (ATP
 pyrophosphate-lyase) (Rutabaga protein).
 RUT.
 Drosophila melanogaster (Fruit fly).
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 [1]_TaxID=7227;
 SEQUENCE FROM N.A.
 STRAIN=Canton-S; TISSUE=Head;
 MEDLINE=92154664; PubMed=173965;
 Levin L.R., Han P.-L., Kwang P.M., Feinstein P.G., Davis R.L.,
 Reed R.R.;
 "The Drosophila learning and memory gene rutabaga encodes a
 Ca2+/calmodulin-responsive adenylyl cyclase.";
 Cell 68:479-489(1992).
 CC -!- FUNCTION: This is a membrane-bound, calmodulin-sensitive adenylyl
 cyclase. Inactivation of this cyclase leads to a learning and
 memory defect.
 CC -!- CATALYTIC ACTIVITY: ATP = 3',5'-cyclic AMP + diphosphate.
 CC -!- COFACTOR: Binds 2 magnesium ions per subunit (by similarity).
 CC -!- ENZYME REGULATION: Activated by calcium/calmodulin and G protein.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- TISSUE SPECIFICITY: MUSHROOM BODIES OF THE FLY BRAIN.
 CC -!- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.
 CC -!- SIMILARITY: Belongs to the adenylyl cyclase class-4/guanylyl
 cyclase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M81887; AAA28844.1; -.
 CC F01; D42088; D42088.
 CC HSP; P19754; 1AWK.
 CC FlyBase; Fgn0003301; rut.
 CC GO; GO:0019933; P: cAMP-mediated signaling; NAS.
 CC GO; GO:0007625; P: grooming behavior; NAS.
 CC GO; GO:0007591; P: molting cycle (sensu Insecta); IGI.
 CC GO; GO:0008355; P: olfactory learning; NAS.
 CC GO; GO:0045473; P: response to ethanol (sensu Insecta); NAS.
 CC InterPro; IPR001054; G_cyclase
 CC Pfam; PF00211; Guanylate_cyc; 2.
 CC SMART; SM00044; CYCC; 2.
 CC PROSITE; PS00452; GUANYLATE_CYCLASES_1; 2.
 CC PROSITE; PS0125; GUANYLATE_CYCLASES_2; 2.
 CC Lyase; CAMP biosynthesis; Transmembrane; Glycoprotein; Repeat;
 CC Metal-binding; Magnesium.
 CC DOMAIN 1 41 CYTOPLASMIC (POTENTIAL).
 CC TRANSMEM 42 60 POTENTIAL.
 CC TRANSMEM 65 84 POTENTIAL.
 CC TRANSMEM 101 115 POTENTIAL.
 CC TRANSMEM 122 142 POTENTIAL.
 CC TRANSMEM 152 174 POTENTIAL.
 CC TRANSMEM 186 206 POTENTIAL.
 CC DOMAIN 207 705 CYTOPLASMIC (POTENTIAL).
 CC TRANSMEM 706 726 POTENTIAL.
 CC TRANSMEM 730 750 POTENTIAL.
 CC TRANSMEM 770 791 POTENTIAL.
 CC DOMAIN 792 813 EXTRACELLULAR (POTENTIAL).
 CC TRANSMEM 814 834 POTENTIAL.
 CC TRANSMEM 842 867 POTENTIAL.
 CC TRANSMEM 868 888 POTENTIAL.
 CC DOMAIN 889 2248 CYTOPLASMIC (POTENTIAL).
 CC TRANSMEM 318 334 CATALYTIC (POTENTIAL).
 CC DOMAIN 1013 1029

FT DOMAIN 515 530 GLY-RICH.
 FT DOMAIN 569 602 GLY-RICH.
 FT DOMAIN 1278 1297 GLN-RICH.
 FT DOMAIN 1767 1810 GLY/SER-RICH.
 FT DOMAIN 2025 2040 ASP/GLU-RICH (ACIDIC).
 FT DOMAIN 2200 2241 GLN-RICH.
 FT METAL 280 MAGNESIUM 1 AND 2 (BY SIMILARITY).
 FT METAL 281 MAGNESIUM 2 (VIA CARBONYL OXYGEN) (BY
 SIMILARITY).
 FT METAL 324 324 MAGNESIUM 1 AND 2 (BY SIMILARITY).
 FT CARBOHYD 800 800 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 807 807 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT MUTAGEN 1026 1026 G->R: ABOLISHES CATALYTIC ACTIVITY.
 SQ SEQUENCE 2248 AA; 248899 MW; E459C718BE018868 CRC64;
 Query Match 73.9%; Score 34; DB 1; Length 2248;
 Best Local Similarity 75.0%; Pred. No. 3.3e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 STEGEGSG 8
 Db 1598 TTEGEGEG 1605
 ID SNP2 STRCO STANDARD; PRT; 227 AA.
 AC P43152;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Extracellular small neutral protease precursor (EC 3.4.24.77)
 DE (Snapsalysin) (Extracellular metalloprotease).
 GN SNPA OR MPRA2 OR PRT OR LMP OR SC07432 OR SC6D11.28C.
 OS Streptomyces coelicolor, and
 OS Streptomyces lividans.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 CC NCBI_TaxID=1902; 1916;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.lividans; STRAIN=66 / 1326;
 RX MEDLINE=93099553; PubMed=1464066;
 RA Butler M.J., Davey C.C., Krygman P., Walczyk E., Malek L.T.;
 RT "Cloning of genetic loci involved in endoprotease activity in
 RT Streptomyces lividans 66: a novel neutral protease gene with an
 RT adjacent divergent putative regulatory gene.";
 RL Can. J. Microbiol. 38:912-920(1992).
 RN [2]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 113-131.
 RC SPECIES=S.lividans; STRAIN=TK24;
 RX MEDLINE=92192468; PubMed=1547948;
 RA Lichenstein H.S., Busse L.A., Smith G.A., Nishi L.O.,
 RA McInley M.O., Rohde M.F., Katzwitz J.L., Zukowski M.M.;
 RT "Cloning and characterization of a gene encoding extracellular
 RT metalloprotease from Streptomyces lividans.";
 RL Gene 111:125-130(1992).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.lividans;
 RX Takahashi H.;
 RL Submitted (SEP-1992) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.coelicolor; STRAIN=A3(2) / M145;
 RX MEDLINE=21996410; PubMed=12000953;
 RA Thomson S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
 RA Bentley N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
 RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares S., Taylor K.,

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 STEGSGR 9
 DB 4 NTEAQSQR 12

RESULT 6
 TSY_YEAST
 ID TSY_YEAST STANDARD; PRT; 304 AA.
 AC P06785; Q12694;
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DE Thymidylate synthase (EC 2.1.1.45) (TSase).
 GN TMPI OR CD21 OR YOR074C OR YOR29-25
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 NCBI_TaxID=4932;
 [1]
 RN
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87165970; PubMed=3031048;
 RA Taylor G.R., Lagosky P.A., Storms R.K., Haynes R.H.;
 RT "Molecular characterization of the cell cycle-regulated thymidylate
 RT synthase gene of Saccharomyces cerevisiae."
 RL J. Biol. Chem. 262:5298-5307 (1987).
 [2]
 RN
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97279235; PubMed=9133743;
 RA Valens M., Bohn C., Daignan-Pornier B., Dang V., Bolotin-Fukuhara M.;
 RT "The presence of a 54.7 Kb fragment of yeast chromosome XV reveals
 RT the presence of two CRNAs and 24 new open reading frames."
 RL Yeast 13:379-390 (1997).
 [3]
 RN
 RP SEQUENCE OF 1-38 FROM N.A.
 RX MEDLINE=89096830; PubMed=3062362;
 RA McIntosh E.M., Ord R.W., Storms R.K.;
 RT "Transcriptional regulation of the cell cycle-dependent thymidylate
 RT synthase gene of Saccharomyces cerevisiae."
 RL Mol. Cell. Biol. 9:4616-4624 (1989).
 CC -1- FUNCTION: REQUIRED FOR BOTH NUCLEAR AND MITOCHONDRIAL DNA
 CC SYNTHESIS.
 CC -1- CATALYTIC ACTIVITY: 5,10-methylenetetrahydrofolate + dUMP =
 CC dihydrofolate + dTMP
 CC -1- PATHWAY: Deoxyribonucleotide biosynthesis.
 CC -1- SUBUNIT: Homodimer.
 CC -1- SIMILARITY: BELONGS TO THE THYMIDYLATE SYNTHASE FAMILY.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).

EMBL; J02706; AAA60940.1; --
 DR EMBL; Z74982; CAA9267.1; ALT_SEQ.
 DR EMBL; Z70678; CAA94559.1; --
 DR EMBL; M29100; AAA35159.1; --
 DR PIR; S66957; YXBYT.
 DR HSP; F04818; IHW4.
 DR SGD; S0005600; CDC21.
 DR GO; GO:0005634; C:nucleus; IDA.
 DR GO; GO:0004799; F:thymidylate synthase activity; IDA.
 DR InterPro; IPR000398; Thymidylat_synth.
 DR Pfam; PF00303; thymidylat_synth; 1.
 DR PRINTS; PR00108; THYMSYNTHASE.
 DR PRODOM; PD001180; Thymidylat_synth; 1.
 DR PROSITE; PS00091; THYMIDYLATE SYNTHASE; 1.
 DR Transferase; Methyltransferase; Nucleotide biosynthesis.
 ACT_SITE 177 177
 BY SIMILARITY.

SQ SEQUENCE 304 AA; 35047 MW; 0C514BEDB8574510 CRC64;

Query Match 73.9%; Score 34; DB 1; Length 304;
 Best Local Similarity 85.7%; Pred. No. 40;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 EGEQSQR 9
 DB 190 EGEQSQR 196

RESULT 7
 BEXD_HAEIN
 ID BEXD_HAEIN STANDARD; PRT; 394 AA.
 AC P22236;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Capsule polysaccharide export protein bexD precursor.
 GN BEXD.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Haemophilus.
 NCBI_TaxID=727;
 [1]
 RN
 RP SEQUENCE FROM N.A.
 RX STRAIN=Eagan / Serotype B;
 RX MEDLINE=91186821; PubMed=2082145;
 RA Kroll J.S., Loynds B., Brophy L.N., Moxon E.R.;
 RT "The bex locus in encapsulated Haemophilus influenzae: a chromosomal
 RT region involved in capsule polysaccharide export."
 RL Mol. Microbiol. 4:1853-1862 (1990).
 CC -1- FUNCTION: May form an ATP-driven capsule polysaccharide export
 CC apparatus, in association with the bexA, bexB and bexC proteins.
 CC May function as a membrane anchor for capsular polysaccharides.
 CC Possible porin properties.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane
 CC (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE BEXD/CTRA/VEXA FAMILY.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).

EMBL; X54987; CAA38730.1; --
 DR PIR; S12232; BMHXD.
 DR InterPro; IPR003715; Poly_export.
 DR Pfam; PF02563; Poly_export; 1.
 DR PROSITE; PS00013; PROKAR LIPOPROTEIN; 1.
 DR Transport; Polysaccharide transport; Porin; Bacterial capsule;
 KW Outer membrane; Signal; Transmembrane; Lipoprotein.
 FT SIGNAL 1 19
 FT CHAIN 20 394
 FT CAPSULE POLYSACCHARIDE EXPORT PROTEIN
 FT BEXD.
 FT LIPID 20 20
 FT N-ACYL DIGLYCERIDE (POTENTIAL).
 SQ SEQUENCE 394 AA; 41973 MW; EDEE230B4D9E4C08 CRC64;

Query Match 73.9%; Score 34; DB 1; Length 394;
 Best Local Similarity 75.0%; Pred. No. 52;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 STEGSGR 8
 DB 111 SSEGQSGS 118

RESULT 8
 CPT2_BRAOL
 ID CPT2_BRAOL STANDARD; PRT; 402 AA.

```
FT DOMAIN 985 1021 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 1022 1042 13 (BY SIMILARITY).
FT DOMAIN 1043 1085 CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM 1086 1106 14 (BY SIMILARITY).
FT DOMAIN 1107 1107 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 1108 1128 15 (BY SIMILARITY).
FT DOMAIN 1129 1199 CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM 1200 1220 16 (BY SIMILARITY).
FT DOMAIN 1221 1222 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 1223 1243 17 (BY SIMILARITY).
FT DOMAIN 1244 1527 CYTOPLASMIC (BY SIMILARITY).
FT NP_BIND 661 668 ATP (POTENTIAL).
FT NP_BIND 1323 1330 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 18 18 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1006 1006 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1007 1007 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARSPLIC 226 510 MAIYGRHPLKEDKLSWSEDRSMVQQLERWRRKQEKQ
TARHKASAPGNAGSEDEVILGAPRPRKSEFKALLATF
GSSFLISACFKLIQDLISFIPQLLSILIRFISNFMPSMW
GFLVAGLMLCSMMQSLIOLHYHYFVTVGVKFTGIMGVI
YKALQIVNSVKRSTVGEIVNLMGVDAQRFMDLAPFLNLL
WSAPLQIILAIYFLQNGLSPVLGAVFWMILLIPLNGAVV
KWEAFQVKMKLDSRIKIMSEILLNGIKVLIYAKESF
-> LNPDPRLRCGLPFTSPDCHLNLPAFSGGEGPLVPKG
RGCIIDPGAAAGGMEERAGADGTTQGFSSWEKLRKRG
AAGCAQAPAEALLPEGPAGHLRLPHQCLLIQVPGPALH
QSTAAQHFDQVYLOPHGFLVLPGWADVPVLHDAVDLIT
TLLPLHLCDWGVSYVDHCHQEGSLQSGVHLSQTCVHCQGN
COPHVGCCPALMGPCPLQCAVAVSTPADHPGDLPLAEPRS
LPGWRSRFLGLADSTQSCGRDADALPGKNAIEGLAHQAD
E (in isoform 3B).
/FTID=VSP.000040.
Missing (in isoform 3B).
/FTID=VSP.000041.
WLSIGVEFVGNVCVLFPAALFAVIGRSSLNPGLVGLSYSL
QVTFALNMIRMSDLGNIVAVRVKYSKTETEARPVVE
GSRPEGPPEPRGEFRNYSVYRPLGLDLVLRDLSLVHGG
EKVGIVGTGAGKSSNTLCFLFLLAAKGEINRIDLNVADI

Query Match 76.1%; Score 35; DB 1; Length 1527;
Best Local Similarity 66.7%; Pred. No. 1.4e-02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 STEGSGQR 9
Db 911 SSGEGQGR 919

RESULT 4
SECE_ECOLI STANDARD; PRT; 127 AA.
AC P16970;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Preprotein translocase secE subunit.
GN SECE OR PRLG OR B3981 OR C4936 OR Z5554 OR ECS4904.
OS Escherichia coli,
OS Escherichia coli O6, and
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562, 217992, 83334;
[1]
SEQUENCE FROM N.A.
RX MEDLINE=90170882; PubMed=2137819;
RA Downing W.L., Sullivan S.L., Gottesman M.E., Dennis P.P.;
RT "Sequence and transcriptional pattern of the essential Escherichia coli secE-nusG operon.";
RL J. Bacteriol. 172:1621-1627(1990).
[2]
SEQUENCE FROM N.A.
RX MEDLINE=89378734; PubMed=2673920;
RA Flower A.M., Osborne R.S., Silhavy T.J.;
```

```
RA Schatz P.J., Riggs P.D., Jacq A., Fath M.J., Beckwith J.;
RT "The secE gene encodes an integral membrane protein required for protein export in Escherichia coli.";
RL Genes Dev. 3:1035-1044(1989).
[3]
SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=94089392; PubMed=8265357;
RA Blattner F.R., Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L.;
RT "Analysis of the Escherichia coli genome. IV. DNA sequence of the region from 89.2 to 92.8 minutes.";
RL Nucleic Acids Res. 21:5408-5417(1993).
[4]
SEQUENCE FROM N.A.
RC STRAIN=06:HI / CFT073 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P., Rasko D., Buckles B.L., Liou S.-R., Boutin A., Hackett J., Stroud D., Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T., Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
[5]
SEQUENCE FROM N.A.
RC STRAIN=0157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D., Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A., Posfai E., Hackett J., Klink S., Boutin A., Shao Y., Miller L., Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoudis K., Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C., Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
[6]
SEQUENCE FROM N.A.
RC STRAIN=0157:H7 / RMD 050952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Maxino K., Onishi K., Kurokawa K., Ishii K., Yokoyama K., Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T., Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T., Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
[7]
TOPOLOGY.
RC MEDLINE=91266903; PubMed=2050112;
RA Schatz P.J., Bieker K.L., Ottmann K.M., Silhavy T.J., Beckwith J.;
RT "One of three transmembrane stretches is sufficient for the functioning of the SecE protein, a membrane component of the E. coli secretion machinery.";
RL EMO J. 10:1749-1757(1991).
[8]
SUBUNITS.
RX MEDLINE=20402345; PubMed=10944122;
RA Yahr T.L., Wickner W.T.;
RT "Evaluating the oligomeric state of SecYEG in preprotein translocase.";
RL EMO J. 19:4393-4401(2000).
[9]
SUBUNITS.
RX MEDLINE=20164453; PubMed=10698927;
RA Manting E.H., van Der Does C., Remigy H., Engel A., Driessen A.J.;
RT "SecYEG assembles into a tetramer to form the active protein translocation channel.";
RL EMO J. 19:852-861(2000).
[10]
MUTANTS.
RX MEDLINE=95196752; PubMed=7889938;
RA Flower A.M., Osborne R.S., Silhavy T.J.;
```

```

DR PROSITE; PS50885; HAMP; 1.
DR PROSITE; PS50109; HIS_KIN; 1.
DR PROSITE; PS50112; PAS; FALSE_NEG.
KW Sensory transduction; Transfrase; Kinase; Phosphorylation;
KW Nitrogen fixation; Transmembrane.
FT TRANSMEM 29 49 POTENTIAL.
FT TRANSMEM 61 81 POTENTIAL.
FT TRANSMEM 107 127 POTENTIAL.
FT TRANSMEM 202 222 POTENTIAL.
FT TRANSMEM 305 325 POTENTIAL.
FT DOMAIN 327 380 HAMP.
FT DOMAIN 392 465 PAS.
FT DOMAIN 523 745 HISTIDINE KINASE.
SQ SEQUENCE 777 AA; 84353 MW; 23C524D2F0631D76 CRC64;

Query Match 78.3%; Score 36; DB 1; Length 777;
Best Local Similarity 87.5%; Pred. No. 46;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 STEGEGSG 8
Db 483 STEGEGSG 490

RESULT 2
STAU_DROME STANDARD; PRT; 1026 AA.
AC P25159;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Maternal effect protein stauften.
GN STAU.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN NCBI_TaxID=7227;
RP SEQUENCE FROM N.A.
RX MEDLINE=91300532; PubMed=1712672;
RA St Johnston D., Beuchle D., Nusslein-Volhard C.;
RT "Staufen, a gene required to localize maternal RNAs in the Drosophila
RT egg.";
RL Cell 66:51-63(1991).
RN [2]
RP STRUCTURE BY NMR OF 579-646.
RX MEDLINE=95354674; PubMed=7628456;
RA Bycroft M., Grunert S., Murzin A.G., Proctor M., St Johnston D.;
RT "NMR solution structure of a dsRNA binding domain from Drosophila
RT stauften protein reveals homology to the N-terminal domain of
RT ribosomal protein S5.";
RL EMBO J. 14:3563-3571(1995).
RN [3]
RP ERRATUM.
RA Bycroft M., Grunert S., Murzin A.G., Proctor M., St Johnston D.;
RL EMBO J. 14:4385-4385(1995).
RN [4]
RP CHARACTERIZATION OF DBRM DOMAINS.
RX MEDLINE=20183617; PubMed=10716936;
RA Mcklem D.R., Adams J., Grunert S., St Johnston D.;
RT "Distinct roles of two conserved Stauften domains in oskar mRNA
RT localization and translation.";
RL EMBO J. 19:1368-1377(2000).
CC -1- FUNCTION: REQUIRED BOTH FOR THE LOCALIZATION OF MATERNAL
CC DETERMINANTS TO THE POSTERIOR POLE OF THE DROSOPHILA EGG AND
CC FOR BICOID RNA TO LOCALIZE CORRECTLY TO THE ANTERIOR POLE.
CC OSKAR PROTEIN IS REQUIRED TO KEEP OSKAR RNA AND STAUFEN PROTEIN AT
CC THE POSTERIOR POLE.
CC -1- TISSUE SPECIFICITY: POLAR GRANULES AT THE POSTERIOR POLE OF THE
CC OOCTYE, AND BY THE TIME THE EGG IS LAID, AT THE ANTERIOR POLE.
CC -1- DOMAIN: CONTAINS A PROLINE-RICH DOMAIN. THE INSERTION OF THIS
CC DOMAIN IN THE DBRM 2 DOMAIN IS REQUIRED FOR STAU-OSKAR MENA

LOCALIZATION
-1- SIMILARITY: Contains 5 DBRM (double-stranded RNA-binding) domains.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
EMBL; M69111; AAA73062.1; -
DR PIR; A40315; A40315.
DR PDB; 1STU; 31-JUL-95.
DR PDB; 1EKZ; 21-AUG-00.
DR FlyBase; FBGN0003520; stau.
DR GO; GO:0045179; C:apical cortex; IDA.
DR GO; GO:0003730; F:mRNA 3' UTR binding activity; IDA.
DR GO; GO:0045034; F:neuroblast cell division; NAS.
DR GO; GO:0045450; P:pole plasm bicoid mRNA localization; NAS.
DR GO; GO:0007316; P:pole plasm RNA localization; IMP.
DR InterPro; IPR01159; DS_RBD.
DR Pfam; PF00035; dsrm; 5.
DR SMART; SM00358; DSRM; 4.
DR PROSITE; PS50137; DS_RBD; 5.
KW Developmental protein; RNA-binding; Repeat; 3D-structure.
FT DOMAIN 311 378 DBRM 1.
FT DOMAIN 398 557 DBRM 2 (ATYPICAL).
FT DOMAIN 578 645 DBRM 3.
FT DOMAIN 711 781 DBRM 4.
FT DOMAIN 951 1018 DBRM 5.
FT DOMAIN 3 77 GLN/HIS/PRO-RICH.
FT DOMAIN 42 46 POLY-GLN.
FT DOMAIN 47 51 POLY-PRO.
FT DOMAIN 71 77 POLY-GLN.
FT DOMAIN 451 455 POLY-PRO.
FT HELIX 580 590
FT TURN 591 591
FT TURN 595 598
FT STRAND 609 616
FT TURN 617 618
FT STRAND 619 627
FT TURN 628 629
FT HELIX 630 644
SQ SEQUENCE 1026 AA; 110347 MW; A55B97624BBF7D0B CRC64;

Query Match 78.3%; Score 36; DB 1; Length 1026;
Best Local Similarity 75.0%; Pred. No. 62;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TELEGSGR 9
Db 621 TELEGNGK 528

RESULT 3
MRP3 HUMAN STANDARD; PRT; 1527 AA.
AC O15438; O60265; O60922; O75621; O95078; O95289; O95290; Q9UN52;
DT 15-JUL-1998 (Rel. 36, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Canalicular multispecific organic anion transporter 2 (Multidrug
DE resistance-associated protein 3) (Multi-specific organic anion
DE transporter-2) (MOAT-D).
GN ABCG3 OR CMOAT2 OR MRP3 OR MRP2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN NCBI_TaxID=9606;
RP SEQUENCE FROM N.A.
RX MEDLINE=98409330; PubMed=9738950;

```


A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AC0934
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-127 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD09493.1; PID:G16504610; GSPDB:GN00176
C:Genetics:
A:Gene: STV3738
C:Superfamily: protein-export protein secE

Query Match 73.9%; Score 34; DB 2; Length 127;
Best Local Similarity 66.7%; Pred. No. 36;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 STEGSGR 9
||:||||
Db 4 NTEAGSGR 12

RESULT 12
AC3193
A:Title: Glutathione S-transferase Atu5272 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
A:Accession: AC3193
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.; Erage, G.; Gilliet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClellan, J.; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AC3193
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-203 <KUR>
A:Cross-references: GB:AE008687; PIDN:AAL45961.1; PID:G17743713; GSPDB:GN00188
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atus272
A:Genome: plasmid
C:Superfamily: hypothetical protein b0838

Query Match 73.9%; Score 34; DB 2; Length 203;
Best Local Similarity 75.0%; Pred. No. 58;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 TEGSGR 9
||:||||
Db 36 TRDGSGR 43

RESULT 13
B83418
A:Title: DNA polymerase III, epsilon chain PA1816 [imported] - Pseudomonas aeruginosa (strain PAC C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: B83418
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Buchanan, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Latbig, K.; Lim, J.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
A:Reference number: AB2950; MUID:20437337; PMID:10984043
A:Accession: B83418
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-246 <STO>
A:Cross-references: GB:AE004608; GB:AE004091; NID:G9947797; PIDN:AAG05205.1; GSPDB:GN00189
A:Experimental source: strain PA01
C:Genetics:
A:Gene: dnaQ; PA1816

C:Superfamily: dnaQ protein

Query Match 73.9%; Score 34; DB 2; Length 246;
Best Local Similarity 85.7%; Pred. No. 70;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 EGEGR 9
||:||||
Db 185 EGDGSGR 191

RESULT 14
YXBT
A:Title: thymidylate synthase (EC 2.1.1.45) - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein O2950; protein YOR074c
C:Species: Saccharomyces cerevisiae
C:Date: 31-Mar-1993 #sequence_revision 13-Mar-1997 #text_change 16-Jun-2000
A:Accession: S66957; A29546
R:Bohn, C.; Bolotin-Fukuhara, M.; Daignan-Fornier, B.; Dang, D.V.; Valens, M.
submitted to the Protein Sequence Database, July 1996
A:Reference number: S66929
A:Accession: S66957
A:Molecule type: DNA
A:Residues: 1-259 <BOH>
A:Cross-references: ENBL:Z74982; NID:G1420228; PIDN:CAA99267.1; PID:G1420229; GSPDB:GN00189
A:Experimental source: strain S288C
R:Taylor, G.R.; Lagosky, P.A.; Storms, R.K.; Haynes, R.H.
J. Biol. Chem. 262, 5298-5307, 1987
A:Title: Molecular characterization of the cell cycle-regulated thymidylate synthase gene
A:Reference number: A29546; MUID:87165970; PMID:3031048
A:Accession: A29546
A:Molecule type: DNA
A:Residues: 1-33, 'GTSLFAPPQLRSLRDTFLLTTKVFTRGIIELWFLAGDT', 34-259 <TAY>
A:Cross-references: GB:J02706; NID:G172989; PIDN:AAA60940.1; PID:G172990
C:Genetics:
A:Gene: SGD:CDC21; TMPI, MIPS:YOR074c
A:Cross-references: SGD:S0005600; MIPS:YOR074c
A:Map position: 15R
A:Introns: 34/1
C:Superfamily: thymidylate synthase; thymidylate synthase homology
C:Keywords: deoxyribonucleotide biosynthesis; homodimer; methyltransferase; nucleus; pyrophosphate; domain: thymidylate synthase homology <IDS>
F:10-259/Domain: thymidylate synthase homology <IDS>
F:332/Active site: Cys #status predicted

Query Match 73.9%; Score 34; DB 1; Length 259;
Best Local Similarity 85.7%; Pred. No. 74;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 EGEGR 9
||:||||
Db 145 EGEGR 151

RESULT 15
AI2079
A:Title: hypothetical protein alr2191 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C:Accession: AI2079
R:Kaneke, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, N.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, N.
DNA Res. 9, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AI2079
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-305 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA073690.1; PID:G17131282; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr2191


```

QY      1 STEGSGR 9
Db      132 STYEGSGR 140

RESULT 3
T04103
sucrose-phosphate synthase (EC 2.4.1.14) 1 - rice
C;Species: Oryza sativa (rice)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 20-Jun-2000
C;Accession: T04103
R;Sakamoto, M.; Satozawa, T.; Kishimoto, N.; Higo, K.; Shimada, H.; Fujimura, T.
Plant Sci. 112, 207-217, 1995
A;Title: Structure and RFLP mapping of a rice sucrose phosphate synthase (SPS) gene that
A;Reference number: Z15212
A;Accession: T04103
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1084 <SAK>
A;Cross-references: EMBL:D45890; PIDN:PA08304.1
A;Experimental source: subsp. Japonica
C;Genetics:
A;Gene: Sps1
A;Map position: 1
A;Intons: 120/3; 200/2; 221/3; 452/3; 536/3; 613/3; 634/3; 946/3; 989/2
C;Superfamily: sucrose-phosphate synthase; sucrose/sucrose-phosphate synthase homology
C;Keywords: Glycosyltransferase; hexosyltransferase
F;196-680/Domain: sucrose/sucrose-phosphate synthase homology <SSPS>

Query Match      82.6%; Score 38; DB 2; Length 1084;
Best Local Similarity 87.5%; Pred. No. 58;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 STEGSGR 8
Db      268 STDGSGR 275

RESULT 4
F82581
ribonuclease III XP2246 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
C;Accession: F82581
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: F82581
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-212 <SIM>
A;Cross-references: GB:AE003849; NID:g9107394; PIDN:AAF85045.1; GSPDB:GN001
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Rehnach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.B.A.; Carraro, D.M.; Carrier, H
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tshukako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XP2246
C;Superfamily: ribonuclease III; double-stranded RNA-binding repeat homology

Query Match      78.3%; Score 36; DB 2; Length 212;
Best Local Similarity 87.5%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 TEEGSGR 9
Db      186 TEEGSGR 193

RESULT 5
T34704
probable transferase - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999
C;Accession: T34704
R;Oliver, K.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, May 1998
A;Reference number: Z21554
A;Accession: T34704
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-697 <OLI>
A;Cross-references: EMBL:AL023702; PIDN:CAA19236.1; GSPDB:GN00070; SCOEDB:SC1C3.12
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SC1C3.12

Query Match      78.3%; Score 36; DB 2; Length 697;
Best Local Similarity 77.8%; Pred. No. 86;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 STEGSGR 9
Db      425 STEPGDGR 433

RESULT 6
A40315
maternal effect protein (staufen) - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A40315
R;St. Johnston, D.; Beuchle, D.; Nusslein-Volhard, C.
Cell 66, 51-63, 1991
A;Title: staufen, a gene required to localize maternal RNAs in the Drosophila egg.
A;Reference number: A40315; MUID:91300552; PMID:1712672
A;Accession: A40315
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1026 <STJ>
A;Cross-references: GB:M69111; NID:gi58505; PIDN:AAA73062.1; PID:gi58506
C;Genetics:
A;Gene: FlyBase:stau
A;Cross-references: FlyBase:FBgn0003520
C;Superfamily: maternal effect protein; double-stranded RNA-binding repeat homology
F;308-379/Domain: double-stranded RNA-binding repeat homology <DSR1>
F;575-546/Domain: double-stranded RNA-binding repeat homology <DSR2>
F;708-782/Domain: double-stranded RNA-binding repeat homology <DSR3>

Query Match      78.3%; Score 36; DB 1; Length 1026;
Best Local Similarity 75.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      2 TEEGSGR 9
Db      621 TELEGNGK 628

RESULT 7
JE0336
canalicular multispecific organic anion transporter - human
C;Species: Homo sapiens (man)

```

Fri Dec 12 13:22:39 2003

```

; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C56
; CURRENT APPLICATION NUMBER: US/10/243,261
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 54
; LENGTH: 839
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-243-261-54

```

```

Query Match      73.9%; Score 34; DB 12; Length 839;
Best Local Similarity 85.7%; Pred. No. 1.3e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      3 EGEGR 9
      |||:|
Db      818 EGEGR 824

```

Search completed: December 11, 2003, 18:38:03
Job time : 50.3333 secs

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 EGEGR 9
Db 226 EGEGR 232

RESULT 10

US-09-738-626-6175
; Sequence 6175, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: Patent in ver. 3.0
; SEQ ID NO 6175
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6175

Query Match 73.9%; Score 34; DB 10; Length 411;
Best Local Similarity 66.7%; Pred. No. 6.8e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 STEGR 9
Db 285 NTDGGLR 293

RESULT 11

US-10-237-496-54
; Sequence 54, Application US/10237496
; Publication No. US20030138896A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Batton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: F3630R1C4
; CURRENT APPLICATION NUMBER: US/10/237,496
; CURRENT FILING DATE: 2002-09-06
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18

; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 54
; LENGTH: 839
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-237-496-54

Query Match 73.9%; Score 34; DB 12; Length 839;
Best Local Similarity 85.7%; Pred. No. 1.3e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 EGEGR 9
Db 818 EGEGR 824

RESULT 12

US-10-242-074-54
; Sequence 54, Application US/10242074
; Publication No. US20030138897A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Batton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: F3630R1C19
; CURRENT APPLICATION NUMBER: US/10/242,074
; CURRENT FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18

; APPLICANT: Creelman, Robert
 ; APPLICANT: Dubell, Arnold
 ; APPLICANT: Heard, Jacqueline
 ; APPLICANT: Jiang, Cai-Zhong
 ; APPLICANT: Keddie, James
 ; APPLICANT: Pilgrim, Marsha
 ; APPLICANT: Ratcliffe, Oliver
 ; APPLICANT: Reuber, Lynne
 ; APPLICANT: Riechmann, Jose Luis
 ; APPLICANT: Yu, Guo-Liang
 ; APPLICANT: Pinada, Omalra
 ; TITLE OF INVENTION: Genes for Modifying Plant Traits IV
 ; FILE REFERENCE: MBI-0025
 ; CURRENT APPLICATION NUMBER: US/09/934,455
 ; CURRENT FILING DATE: 2001-08-22
 ; PRIOR APPLICATION NUMBER: 60/227439
 ; PRIOR FILING DATE: 2000-08-22
 ; PRIOR APPLICATION NUMBER: MBI-0022
 ; PRIOR FILING DATE: 2001-11-16
 ; PRIOR APPLICATION NUMBER: MBI-0023
 ; PRIOR FILING DATE: 2001-04-17
 ; NUMBER OF SEQ ID NOS: 516
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 38
 ; LENGTH: 174
 ; TYPE: PRT
 ; ORGANISM: Arabidopsis thaliana
 ; US-09-934-455-38

Query Match 84.8%; Score 39; DB 11; Length 174;
 Best Local Similarity 88.9%; Pred. No. 47;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 STEGEGSGR 9
 DB 132 STYEGSGR 140

RESULT 3
 US-10-217-700-9
 ; Sequence 9, Application US/10217700
 ; Publication No. US20030070191A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Haigler, Candace H.
 ; APPLICANT: Holaday, A. Scott
 ; TITLE OF INVENTION: TRANSGENIC FIBER PRODUCING PLANTS WITH INCREASED
 ; TITLE OF INVENTION: EXPRESSION OF SUCROSE PHOSPHATE SYNTHASE
 ; FILE REFERENCE: 201304/1000
 ; CURRENT APPLICATION NUMBER: US/10/217,700
 ; CURRENT FILING DATE: 2002-08-12
 ; EARLIER APPLICATION NUMBER: 09/394,272
 ; EARLIER FILING DATE: 1999-09-10
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 9
 ; LENGTH: 1084
 ; TYPE: PRT
 ; ORGANISM: Oryza sativa
 ; US-10-217-700-9

Query Match 82.6%; Score 38; DB 15; Length 1084;
 Best Local Similarity 87.5%; Pred. No. 3.9e+02;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 STEGEGSG 8
 DB 268 STDGEGSG 275

RESULT 4
 US-10-116-275-5
 ; Sequence 5, Application US/10116275
 ; Publication No. US20030211476A1

; GENERAL INFORMATION:
 ; APPLICANT: Elan Pharmaceutical Technology
 ; APPLICANT: O'Mahony, Daniel J.
 ; APPLICANT: Brayden, David
 ; APPLICANT: Byrne, Daragh
 ; APPLICANT: Lambkin, Imelda
 ; APPLICANT: Higgins, Lisa
 ; TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and
 ; TITLE OF INVENTION: Compositions Targeting Peyer's Patches and M Cell Receptors
 ; FILE REFERENCE: E1067/20087
 ; CURRENT APPLICATION NUMBER: US/10/116,275
 ; CURRENT FILING DATE: 2002-10-04
 ; NUMBER OF SEQ ID NOS: 349
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 5
 ; LENGTH: 38
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Listed in Table titled "Peptides that Target to and/or Enhance t
 ; OTHER INFORMATION: take Across the GIT"
 ; US-10-116-275-5

Query Match 76.1%; Score 35; DB 12; Length 38;
 Best Local Similarity 77.8%; Pred. No. 48;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 STEGEGSGR 9
 DB 1 STPGGSGR 9

RESULT 5
 US-09-815-242-10749
 ; Sequence 10749, Application US/09815242
 ; Patent No. US20020061569A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari L.
 ; APPLICANT: Zyskind, Judith W.
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John D.
 ; APPLICANT: Carr, Grant J.
 ; APPLICANT: Yamamoto, Robert T.
 ; APPLICANT: Xu, H. Howard
 ; TITLE OF INVENTION: Identification of Essential Genes in
 ; TITLE OF INVENTION: Prokaryotes
 ; FILE REFERENCE: ELITRA.011A
 ; CURRENT APPLICATION NUMBER: US/09/815,242
 ; CURRENT FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; NUMBER OF SEQ ID NOS: 14110
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 10749
 ; LENGTH: 609
 ; TYPE: PRT
 ; ORGANISM: Enterococcus faecalis
 ; US-09-815-242-10749

Query Match 76.1%; Score 35; DB 9; Length 609;

Fri Dec 12 13:22:37 2003

; FILING DATE: 12-JUN-1990
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Sweeney, Patricia A.
 ; REGISTRATION NUMBER: 32,733
 ; REFERENCE/DOCKET NUMBER: 0125R2R3
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (515) 248 4897
 ; TELEFAX: (515) 248-4844
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 473 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ;
 US-08-474-404-2

Query Match 73.9%; Score 34; DB 2; Length 473;
 Best Local Similarity 85.7%; Pred. No. 2.8e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 EGEGR 9
 |||||:
 Db 238 EGEGR 244

RESULT 15
 US-08-485-845-2
 ; Sequence 2, Application US/08485845
 ; Patent No. 5850014
 ; GENERAL INFORMATION:
 ; APPLICANT: ALBERTSEN, Marc C.
 ; APPLICANT: BEACH, Larry R.
 ; APPLICANT: HOWARD, John A.
 ; APPLICANT: HUFFMAN, Gary A.
 ; TITLE OF INVENTION: Nucleotide Sequences Mediating Fertility and Method of Using
 ; NUMBER OF SEQUENCES: 8
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pioneer Hi-Bred International, Inc.
 ; STREET: 700 Capital Square, 400 Locust Street
 ; CITY: Des Moines
 ; STATE: Iowa
 ; COUNTRY: U.S.
 ; ZIP: 50309
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/485,845
 ; FILING DATE: 07-JUN-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/103,739
 ; FILING DATE: 02-AUG-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/537,183
 ; FILING DATE: 12-JUN-1990
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Sweeney, Patricia A.
 ; REGISTRATION NUMBER: 32,733
 ; REFERENCE/DOCKET NUMBER: 0125R2
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (515) 248 4897
 ; TELEFAX: (515) 248-4844
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 473 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear

Query Match 73.9%; Score 34; DB 2; Length 473;
 Best Local Similarity 85.7%; Pred. No. 2.8e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 EGEGR 9
 |||||:
 Db 238 EGEGR 244

Search completed: December 11, 2003, 18:30:36
 Job time : 10.2222 secs

Best Local Similarity 77.8%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 STEGSGR 9
Db 1 STPGSGR 9

RESULT 8
US-09-252-991A-20150
; Sequence 20150, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20150
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20150

Query Match 73.9%; Score 34; DB 4; Length 166;
Best Local Similarity 85.7%; Pred. No. 18+02; 0; Indels 0; Gaps 0;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 EGEGR 9
Db 105 EGDGSR 111

RESULT 9
US-09-220-528-116
; Sequence 116, Application US/09220528A
; Patent No. 6284540
; GENERAL INFORMATION:
; APPLICANT: Milbrandt, Jeffrey D.
; FILE REFERENCE: 6029-7998
; CURRENT APPLICATION NUMBER: US/09/220,528A
; CURRENT FILING DATE: 1998-12-24
; EARLIER APPLICATION NUMBER: 09/218,698
; EARLIER FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 60/108,148
; EARLIER FILING DATE: 1998-11-12
; EARLIER APPLICATION NUMBER: 09/163,283
; EARLIER FILING DATE: 1998-09-29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 116
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-220-528-116

Query Match 73.9%; Score 34; DB 3; Length 201;
Best Local Similarity 85.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 EGEGR 9
Db 1 EGEGR 7

RESULT 10
US-09-252-991A-30011
; Sequence 30011, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30011
; LENGTH: 246
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30011

Query Match 73.9%; Score 34; DB 4; Length 246;
Best Local Similarity 85.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 EGEGR 9
Db 185 EGDGSR 191

RESULT 11
US-08-673-814-6
; Sequence 6, Application US/08673814
; Patent No. 6086894
; GENERAL INFORMATION:
; APPLICANT: Inzana, Thomas J.
; TITLE OF INVENTION: RECOMBINANT VACCINE FOR DISEASES CAUSED
; FILE REFERENCE: 6029-7998
; CURRENT APPLICATION NUMBER: US/08/673,814
; CURRENT FILING DATE: 1998-12-24
; EARLIER APPLICATION NUMBER: 09/218,698
; EARLIER FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 60/108,148
; EARLIER FILING DATE: 1998-11-12
; EARLIER APPLICATION NUMBER: 09/163,283
; EARLIER FILING DATE: 1998-09-29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 116
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-673-814-6

Query Match 73.9%; Score 34; DB 3; Length 394;

```
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/280,443
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/197,794
/ FILING DATE: 17-FEB-1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Bak, Mary E.
/ REGISTRATION NUMBER: 31,215
/ REFERENCE/DOCKET NUMBER: WST49AUSA
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 215-540-9206
/ TELEFAX: 215-540-5818
/ INFORMATION FOR SEQ ID NO: 18:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 71 amino acids
/ TYPE: amino acid
/ TOPOLOGY: unknown
/ MOLECULE TYPE: protein
/ US-08-280-443-18

Query Match      78.3%; Score 36; DB 1; Length 71;
Best Local Similarity 75.0%; Pred. No. 21;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      2 TELEGSGR 9
Db      45 TELEGNGK 52

RESULT 3
US-08-457-459-18
; Sequence 18, Application US/08457459
; Patent No. 5677428
; GENERAL INFORMATION:
; APPLICANT: Nishikura, Kazuko
; TITLE OF INVENTION: RNA Editing Enzyme and Methods of Use
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/457,459
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/197,794
; FILING DATE: 17-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/280,443
; FILING DATE: 25-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: WST49CUSA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9206
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 71 amino acids
; TYPE: amino acid
```

```
/ TOPOLOGY: unknown
/ MOLECULE TYPE: protein
/ US-08-457-459-18

Query Match      78.3%; Score 36; DB 1; Length 71;
Best Local Similarity 75.0%; Pred. No. 21;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      2 TELEGSGR 9
Db      45 TELEGNGK 52

RESULT 4
US-08-555-678-18
; Sequence 18, Application US/08555678
; Patent No. 5763174
; GENERAL INFORMATION:
; APPLICANT: Nishikura, Kazuko
; TITLE OF INVENTION: RNA Editing Enzyme and Methods
; TITLE OF INVENTION: of Use Thereof
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, P.O. Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/555,678
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/197,794
; FILING DATE: 17-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/280,443
; FILING DATE: 25-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/457,459
; FILING DATE: 01-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: WST49DUSA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9206
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 71 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
/ US-08-555-678-18

Query Match      78.3%; Score 36; DB 1; Length 71;
Best Local Similarity 75.0%; Pred. No. 21;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      2 TELEGSGR 9
Db      45 TELEGNGK 52

RESULT 5
```

CC	modified trait, the method comprising selecting a polynucleotide that
CC	encodes a polypeptide or an antisense nucleic acid, inserting the
CC	polynucleotide or antisense nucleic acid into an expression vector,
CC	introducing the vector into a plant or a cell of a plant to overexpress
CC	the polypeptide or antisense nucleic acid, thereby producing a modified
CC	plant, and selecting for a modified trait (e.g. increased
CC	production of agriculturally useful proteins or metabolic chemicals,
CC	pest tolerance, environmental stress response (e.g. drought), microbial
CC	disease resistance, herbicide resistance, seed and fruit yield, growth
CC	rate, leaf and flower senescence and many other traits listed in the
CC	specification). The present sequence is one of the 232 proteins which are
CC	A. thaliana transcription factors.
XX	
SQ	Sequence 174 AA;
	Query Match 84.8%; Score 39; DB 23; Length 174;
	Best Local Similarity 88.9%; Pred. No. 35;
	Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0
QY	1 STEGEGSGR 9
DB	132 STYGECSGR 140
RESULT 13	
AAE30378	AAE30378 standard; Protein; 174 AA.
ID	AAE30378
XX	AC AC
XX	AAE30378;
DT	24-FEB-2003 (first entry)
DE	Arabidopsis thaliana G2010 protein.
KW	Phenotype; flowering; floral meristem; plant gene expression regulator;
KW	transcription factor modulator; G2010 protein.
OS	Arabidopsis thaliana.
WO	WO200277185-A2.
PN	03-OCT-2002.
PD	
XX	
FF	26-MAR-2002; 2002WO-US09141.
PR	27-MAR-2001; 2001US-0819142.
PA	(MEND-) MENDEL BIOTECHNOLOGY INC.
PI	Reuber TL;
XX	
DR	WPI; 2003-018910/01.
DR	N-PSDB; AAD47721.
PT	
PT	New isolated or recombinant polynucleotide, useful for modifying the
PT	phenotype of a plant, such as the plant's flowering time or flowering
PT	period leading to commercially superior crops and plants -
XX	
PS	Claim 1; Page 258; 289pp; English.
CC	The present invention relates to recombinant polynucleotides useful for
CC	modifying the phenotype of a plant such as the plant's flowering time
CC	or flowering period leading to commercially superior crops and plants..
CC	Sequences of the invention are useful for modifying the duration of the
CC	phase in which floral meristems are initiated, the duration of time for
CC	which floral organs persist prior to their abscission or the number of
CC	flowers generated on a plant. They can also be used in the recombinant
CC	production of proteins, as regulators of plant gene expression, as
CC	substrates for mutation or PCR reactions, as diagnostic probes for the
CC	presence of complementary or partially complementary nucleic acids or
CC	for identifying exogenous or endogenous modulators of the transcription
CC	factors. The present sequence is Arabidopsis thaliana G2010 protein. This
CC	sequence is used herein exemplification of the invention.


```
PR 02-AUG-1999; 99US-01463388.
PR 02-AUG-1999; 99US-01463389.
PR 03-AUG-1999; 99US-01470308.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-01471192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.

PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 84.8%; Score 39; DB 21; Length 174;
Best Local Similarity 88.9%; Pred. No. 35;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 STEGEGSGR 9
Db 132 STYEGSGR 140

RESULT 11
AAG53081
ID AAG53081 standard; Protein; 174 AA.
XX AC AAG53081;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 67547.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KM hybridisation assay; genetic mapping; gene expression control; promoter;
XX KN termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
```

PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 24-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140891.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 21-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.

PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 04-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158569.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.

PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159337.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 84.8%; Score 39; DB 21; Length 152;
 Best Local Similarity 88.9%; Pred. No. 30;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 STEGEGSGR 9
 DB 110 STYEGSGR 118

RESULT 8
 AAG11547
 ID AAG11547 standard; Protein; 163 AA.
 XX AC AAG11547;
 XX DT 17-OCT-2000 (first entry)
 XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 10304.
 XX KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX OS Arabidopsis thaliana.
 XX PN EP1033405-A2.
 XX PD 06-SEP-2000.
 XX PF 25-FEB-2000; 2000EP-0301439.
 XX PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0121830.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0130891.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 18-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 15-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143824.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.

PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144353.
 PR 20-JUL-1999; 99US-0144354.
 PR 21-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144885.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145090.
 PR 22-JUL-1999; 99US-0145119.
 PR 22-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145228.
 PR 23-JUL-1999; 99US-0145276.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 05-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149829.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.

PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159328.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161922.
 PR 28-OCT-1999; 99US-0161992.
 PR 29-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 84.8%; Score 39; DB 21; Length 152;
 Best Local Similarity 88.9%; Pred. No. 30;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 STEGSGR 9
 |||||
 Db 110 STYEGSGR 118

RESULT 7

AGS3083
 ID AGS3083 standard; Protein; 152 AA.

XX AC AAG53083;

XX DT 18-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 67549.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
 XX hybridisation assay; Genetic mapping; Gene expression control; promoter;
 XX termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

XX PR 05-MAR-1999; 99US-0123180.

XX PR 03-MAR-1999; 99US-0123548.

XX PR 23-MAR-1999; 99US-0125788.

XX PR 25-MAR-1999; 99US-0126264.

XX PR 29-MAR-1999; 99US-0126785.

XX PR 01-APR-1999; 99US-0127462.

XX PR 06-APR-1999; 99US-0128234.

XX PR 08-APR-1999; 99US-0128714.

CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 677 AA;
 Query Match 89.1%; Score 41; DB 22; Length 677;
 Best Local Similarity 100.0%; Pred. No. 62;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 STEGEGSG 8

DB 552 STEGEGSG 559

RESULT 4
 AAY44469
 ID AAY44469 standard; peptide; 9 AA.

AC AAY44469;

XX 27-MAR-2000 (first entry)

DE Human antithrombin III variant Bb.A (residues 385-393).

XX Human; antithrombin III; ATIII variant Bb.A; elastase-resistant;
 KW IgG activated neutrophil resistant; anti-thrombin activity; heparin;
 KW anti-factor Xa activity; blood clotting disorder; sepsis; trauma; stroke;
 KW thrombin activation-related pathological symptom; restenosis; thrombosis;
 KW acute respiratory distress syndrome; thromboembolism; reocclusion.

OS Homo sapiens.
 OS Synthetic.

PH Key Location/Qualifiers

FT Misc-difference 3 /note= "ATIII.N135A Ala at 387 is substituted by Glu"

FT Misc-difference 4 /note= "ATIII.N135A Val at 388 is substituted by Gly"

FT Misc-difference 5 /note= "ATIII.N135A Val at 389 is substituted by Glu"

FT Misc-difference 6 /note= "ATIII.N135A Ile at 390 is substituted by Ala"

FT Misc-difference 7 /note= "ATIII.N135A Ala at 391 is substituted by Ser"

XX WO9958098-A2.

PN 18-NOV-1999.

XX 12-MAY-1999; 99WO-US10549.

XX 12-MAY-1998; 98US-0085197.

XX 05-MAY-1999; 99US-0085197.

PA (BOCK/) BOCK S C.
 PA (PICA/) PICARD V.
 PA (ZEND/) ZENDEHROUH P.

XX Bock SC, Picard V, Zendeirouh P;

XX WPI; 2000-116274/10.

XX New modified human antithrombin III compounds, used for treating e.g.
 PT sepsis, trauma, acute respiratory distress syndrome, restenosis,
 PT thrombosis, thromboembolism or stroke.

XX Claim 13; Page 57; 75pp; English.

XX The present sequence is from an antithrombin III (ATIII) variant, Bb.A
 CC derived from human ATIII.N135A cDNA insert of the pBlueBac baculovirus
 CC expression construct and comprises residues 385-393. The variant has
 CC improved resistance to elastase and IgG-activated neutrophils while
 CC retains anti-thrombin and anti-factor Xa activities. It may be
 CC expressed as glycoforms with enhanced heparin affinity which target the
 CC blood vessel wall more efficiently than ATIII with normal heparin
 CC affinity. The modified ATIII can be used to treat thrombin activation-
 CC related pathological symptoms due to sepsis, trauma, acute
 CC respiratory distress syndrome, restenosis, thrombosis, thromboembolism
 CC and stroke. It can also be used to reduce the risk of reocclusion
 CC and restenosis in percutaneous transluminal coronary angioplasty,
 CC thrombosis associated with surgery, ischaemia/reperfusion injury, and
 CC coagulation abnormalities in cancer or surgical patients.

XX Sequence 9 AA;

Query Match 87.0%; Score 40; DB 21; Length 9;

Best Local Similarity 88.9%; Pred. No. 9.3e+05;

Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 STEGEGSGR 9

DB 1 STEGEGSGR 9

RESULT 5

RAY44472
 ID AAY44472 standard; peptide; 9 AA.

AC AAY44472;

XX 27-MAR-2000 (first entry)

DE Human antithrombin III variant Bb.D (residues 385-393).

XX Human; antithrombin III; ATIII variant Bb.D; elastase-resistant;
 KW IgG activated neutrophil resistant; anti-thrombin activity; heparin;
 KW anti-factor Xa activity; blood clotting disorder; sepsis; trauma; stroke;
 KW thrombin activation-related pathological symptom; restenosis; thrombosis;
 KW acute respiratory distress syndrome; thromboembolism; reocclusion.

OS Homo sapiens.

OS Synthetic.

PH Key Location/Qualifiers

FT Misc-difference 2 /note= "ATIII.N135A Thr at 386 is substituted by Glu"

FT Misc-difference 3 /note= "ATIII.N135A Ala at 387 is substituted by Glu"

FT Misc-difference 4 /note= "ATIII.N135A Val at 388 is substituted by Gly"

FT Misc-difference 5 /note= "ATIII.N135A Val at 389 is substituted by Glu"

FT Misc-difference 6 /note= "ATIII.N135A Ile at 390 is substituted by Gly"

FT Misc-difference 7 /note= "ATIII.N135A Ala at 391 is substituted by Ser"

XX

XX 18-NOV-1999.
 XX 12-MAY-1999; 99WO-US10549.
 XX 12-MAY-1998; 98US-0085197.
 XX 05-MAY-1999; 99US-0085197.
 XX (BOCK/) BOCK S C.
 XX (PICA/) PICARD V.
 XX (ZEND/) ZENDEHROUH P.
 XX Bock SC, Picard V, Zendeherouh P;
 XX WPI; 2000-116274/10.
 XX
 XX New modified human antithrombin III compounds, used for treating e.g.
 XX sepsis, trauma, acute respiratory distress syndrome, restenosis,
 XX thrombosis, thromboembolism or stroke -
 XX
 XX Claim 13; Page 57; 75pp; English.
 XX
 XX The present sequence is from an antithrombin III (ATIII) variant, Bb.B
 XX derived from human ATIII-N135A cDNA insert of the pBlueBac baculovirus
 XX expression construct and comprises residues 385-393. The variant has
 XX improved resistance to elastase and IgG-activated neutrophils while
 XX retains anti-thrombin and anti-factor Xa activities. It may be
 XX expressed as glycoforms with enhanced heparin affinity which target the
 XX blood vessel wall more efficiently than ATIIIs with normal heparin
 XX affinity. The modified ATIIIs can be used to treat thrombin activation-
 XX related pathological symptoms due to sepsis, trauma, acute
 XX respiratory distress syndrome, restenosis, thrombosis, thromboembolism
 XX and stroke. It can also be used to reduce the risk of reocclusion
 XX and restenosis in percutaneous transluminal coronary angioplasty,
 XX thrombosis associated with surgery, ischaemia/reperfusion injury, and
 XX coagulation abnormalities in cancer or surgical patients.
 XX
 XX Sequence 9 AA;
 XX
 XX Query Match 100.0%; Score 46; DB 21; Length 9;
 XX Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX QY 1 STEGEGSGR 9
 XX |||||
 XX Db 1 STEGEGSGR 9
 XX
 XX RESULT 2
 XX ABB14761
 XX ID ABB14761 standard; Protein; 179 AA.
 XX AC ABB14761;
 XX
 XX 23-JAN-2002 (first entry)
 XX
 XX Human nervous system related polypeptide SEQ ID NO 3418.
 XX
 XX Human; neotropic; neuroprotective; cytostatic; dermatological; virucide;
 XX immunosuppressive; anti-inflammatory; anti-HIV; antibacterial; vulnary;
 XX antiparkinsonian; antiskilling; antianaemic; antiarthritic; cancer;
 XX antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
 XX antiallergic; antidiabetic; antilucer; anticonvulsant; antifungal;
 XX antiparasitic; cardiatic; immune disorder; cardiovascular disorder;
 XX neurological disease; infection; nephrotropic; gene therapy; vaccine.
 XX
 XX Homo sapiens.
 XX
 XX WO200159063-A2.
 XX
 XX 16-AUG-2001.
 XX
 XX 17-JAN-2001; 2001WO-US01334.

XX 31-JAN-2000; 2000US-0179065.
 XX 04-FEB-2000; 2000US-0180628.
 XX 24-FEB-2000; 2000US-0184664.
 XX 02-MAR-2000; 2000US-0186350.
 XX 16-MAR-2000; 2000US-0189874.
 XX 17-MAR-2000; 2000US-0190076.
 XX 18-APR-2000; 2000US-0198123.
 XX 19-MAY-2000; 2000US-0205515.
 XX 07-JUN-2000; 2000US-0209467.
 XX 28-JUN-2000; 2000US-0214886.
 XX 30-JUN-2000; 2000US-0215135.
 XX 07-JUL-2000; 2000US-0216647.
 XX 07-JUL-2000; 2000US-0216880.
 XX 11-JUL-2000; 2000US-0217487.
 XX 11-JUL-2000; 2000US-0217496.
 XX 14-JUL-2000; 2000US-0218290.
 XX 26-JUL-2000; 2000US-0220963.
 XX 14-AUG-2000; 2000US-0220964.
 XX 14-AUG-2000; 2000US-0224518.
 XX 14-AUG-2000; 2000US-0224519.
 XX 14-AUG-2000; 2000US-0225213.
 XX 14-AUG-2000; 2000US-0225214.
 XX 14-AUG-2000; 2000US-0225266.
 XX 14-AUG-2000; 2000US-0225267.
 XX 14-AUG-2000; 2000US-0225268.
 XX 14-AUG-2000; 2000US-0225270.
 XX 14-AUG-2000; 2000US-0225447.
 XX 14-AUG-2000; 2000US-0225757.
 XX 14-AUG-2000; 2000US-0225758.
 XX 14-AUG-2000; 2000US-0225759.
 XX 18-AUG-2000; 2000US-0226279.
 XX 22-AUG-2000; 2000US-0226681.
 XX 22-AUG-2000; 2000US-0226868.
 XX 23-AUG-2000; 2000US-0227182.
 XX 23-AUG-2000; 2000US-0227009.
 XX 30-AUG-2000; 2000US-0228924.
 XX 01-SEP-2000; 2000US-0229287.
 XX 01-SEP-2000; 2000US-0229343.
 XX 01-SEP-2000; 2000US-0229344.
 XX 01-SEP-2000; 2000US-0229345.
 XX 05-SEP-2000; 2000US-0229509.
 XX 05-SEP-2000; 2000US-0229513.
 XX 06-SEP-2000; 2000US-0230437.
 XX 06-SEP-2000; 2000US-0230438.
 XX 08-SEP-2000; 2000US-0231242.
 XX 08-SEP-2000; 2000US-0231243.
 XX 08-SEP-2000; 2000US-0231244.
 XX 08-SEP-2000; 2000US-0231413.
 XX 08-SEP-2000; 2000US-0231414.
 XX 08-SEP-2000; 2000US-0232080.
 XX 08-SEP-2000; 2000US-0232081.
 XX 12-SEP-2000; 2000US-0231968.
 XX 14-SEP-2000; 2000US-0232397.
 XX 14-SEP-2000; 2000US-0232398.
 XX 14-SEP-2000; 2000US-0232399.
 XX 14-SEP-2000; 2000US-0232400.
 XX 14-SEP-2000; 2000US-0232401.
 XX 14-SEP-2000; 2000US-0233063.
 XX 14-SEP-2000; 2000US-0233064.
 XX 14-SEP-2000; 2000US-0233065.
 XX 21-SEP-2000; 2000US-0234223.
 XX 21-SEP-2000; 2000US-0234274.
 XX 25-SEP-2000; 2000US-0234997.
 XX 25-SEP-2000; 2000US-0234998.
 XX 26-SEP-2000; 2000US-0235484.
 XX 27-SEP-2000; 2000US-0235834.
 XX 27-SEP-2000; 2000US-0235836.
 XX 29-SEP-2000; 2000US-0236327.
 XX 29-SEP-2000; 2000US-0236367.
 XX 29-SEP-2000; 2000US-0236368.
 XX 29-SEP-2000; 2000US-0236369.
 XX 29-SEP-2000; 2000US-0236370.

```
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NADC-8;
RX MEDLINE=9735197; PubMed=9191863;
RA Andreyev V.G., Wesley R.D., Mengeling W.L., Vorwald A.C., Lager K.M.;
RT "Genetic variation and phylogenetic relationships of 22 porcine
RT reproductively and respiratory syndrome virus (PRRSV) field strains
RT based on sequence analysis of open reading frame 5.";
RL Arch. Virol. 142:993-1001(1997).
DR EMBL; U66394; AAC57968.1; -.
DR InterPro; IPR001332; Arteri_glycop.
DR InterPro; IPR003239; Porcine_RR_virus.
DR Pfam; PF00951; Arteri_glycop; 1.
DR ProDom; PD001151; Porcine_RR_virus; 1.
SQ SEQUENCE 200 AA; 22261 MW; 1D96D1FB91FEC94C CRC64;

Query Match 76.6%; Score 36; DB 12; Length 200;
Best Local Similarity 77.8%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 STEGFFSGR 9
||| ||| |||
Db 97 STAGFFHGR 105

RESULT 10
Q9J7K4 PRELIMINARY; PRT; 200 AA.
AC Q9J7K4
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Envelope glycoprotein.
OS Porcine reproductively and respiratory syndrome virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Arteriviridae; Arterivirus.
OX NCBI_TaxID=28344;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PRRSV19;
RX MEDLINE=20109045; PubMed=10640555;
RA Goldberg T.L., Hahn E.C., Weigel R.M., Scherba G.;
RT "Genetic, geographical and temporal variation of porcine reproductively
RT and respiratory syndrome virus in Illinois.";
RL J. Gen. Virol. 81:171-179(2000).
DR EMBL; AF176441; AAF36247.1; -.
DR InterPro; IPR001332; Arteri_glycop.
DR InterPro; IPR003239; Porcine_RR_virus.
DR Pfam; PF00951; Arteri_glycop; 1.
DR ProDom; PD001151; Porcine_RR_virus; 1.
SQ SEQUENCE 200 AA; 22410 MW; 5DF0D1A59E6ACBB7 CRC64;

Query Match 76.6%; Score 36; DB 12; Length 200;
Best Local Similarity 77.8%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 STEGFFSGR 9
||| ||| |||
Db 97 STAGFFHGR 105

RESULT 11
Q9J7J5 PRELIMINARY; PRT; 200 AA.
AC Q9J7J5
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Envelope glycoprotein.
OS Porcine reproductively and respiratory syndrome virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Arteriviridae; Arterivirus.
OX NCBI_TaxID=28344;
```

```
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PRRSV33;
RX MEDLINE=20109045; PubMed=10640555;
RA Goldberg T.L., Hahn E.C., Weigel R.M., Scherba G.;
RT "Genetic, geographical and temporal variation of porcine reproductively
RT and respiratory syndrome virus in Illinois.";
RL J. Gen. Virol. 81:171-179(2000).
DR EMBL; AF176455; AAF36261.1; -.
DR InterPro; IPR001332; Arteri_glycop.
DR InterPro; IPR003239; Porcine_RR_virus.
DR Pfam; PF00951; Arteri_glycop; 1.
DR ProDom; PD001151; Porcine_RR_virus; 1.
SQ SEQUENCE 200 AA; 22105 MW; 46B84C24AE79D669 CRC64;

Query Match 76.6%; Score 36; DB 12; Length 200;
Best Local Similarity 77.8%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 STEGFFSGR 9
||| ||| |||
Db 97 STAGFFHGR 105

RESULT 12
Q915L2 PRELIMINARY; PRT; 441 AA.
AC Q915L2
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical protein PA0716.
GN PA0716.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warriner P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saler M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
DR EMBL; AB004507; AAG04105.1; -.
DR InterPro; IPR003593; AAA_Atpase.
DR InterPro; IPR003439; ABC_transporter.
DR SMART; SM00382; AAA; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 441 AA; 49872 MW; A8A04B80D49B9F24 CRC64;

Query Match 76.6%; Score 36; DB 16; Length 441;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STEGFFS 7
||| ||| |||
Db 335 STEGFFS 341

RESULT 13
Q81Y79 PRELIMINARY; PRT; 488 AA.
AC Q81Y79
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
```

```
RT from Japan and CWL029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
DR EMBL; AB001686; AAD19194.1; -.
DR EMBL; AB002239; AAF38592.1; -.
DR EMBL; AF002548; BAA99264.1; -.
DR TIGR; CP0793; -.
DR InterPro; IPR004879; DUF255.
DR Pfam; PF03190; DUF255; 1.
KW Complete proteome.
SQ SEQUENCE 700 AA; 80155 MW; 8668C1CFAA17A752 CRC64;

Query Match 80.9%; Score 38; DB 16; Length 700;
Best Local Similarity 77.8%; Pred. No. 35;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 STEGFFSGR 9
Db 367 SREGFFNGR 375

RESULT 2
O55482 PRELIMINARY; PRT; 200 AA.
AC O55482;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Gp5.
OS Porcine reproductive and respiratory syndrome virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Arteriviridae; Arterivirus.
OX NCBI_TaxID=28344;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Danish DK3506-12;
RX MEDLINE=99003922; PubMed=9787654;
RA Madsen K.G.; Hansen C.M.; Madsen E.S.; Strandbygaard B.; Boetner A.;
RA Soerensen K.J.;
RT "Sequence analysis of porcine reproductive and respiratory syndrome
RT virus of the American type collected from Danish swine herds.";
RL Arch. Virol. 143:1683-1700(1998).
DR EMBL; AJ223079; CAAL1088.1; -.
DR InterPro; IPR001332; Arteri_glycop.
DR InterPro; IPR003239; Porcine_RR_virus.
DR Pfam; PF00951; Arteri_glycop; 1.
DR ProDom; PD001151; Porcine_RR_virus; 1.
DR ProDom; PD001151; Porcine_RR_virus; 1.
SQ SEQUENCE 200 AA; 22387 MW; 81295557C68AC3C CRC64;

Query Match 76.6%; Score 36; DB 12; Length 200;
Best Local Similarity 77.8%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 STEGFFSGR 9
Db 97 STAGFFHGR 105

RESULT 3
O41187 PRELIMINARY; PRT; 200 AA.
AC O41187;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Envelope protein.
GN ENV.
OS Porcine reproductive and respiratory syndrome virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Arteriviridae; Arterivirus.
OX NCBI_TaxID=28344;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=41572-2;

from Japan and CWL029 from USA.";
RA Andrejev V.G.; Wesley R.D.; Mengeling W.L.; Vorwald A.C.; Lager K.M.;
RT "Genetic variation and phylogenetic relationships of 22 porcine
RT reproductive and respiratory syndrome virus (PRRSV) field strains
RT based on sequence analysis of open reading frame 5.";
RL Arch. Virol. 142:993-1001(1997).
DR EMBL; U66386; AAC57960.1; -.
DR InterPro; IPR001332; Arteri_glycop.
DR InterPro; IPR003239; Porcine_RR_virus.
DR Pfam; PF00951; Arteri_glycop; 1.
DR ProDom; PD001151; Porcine_RR_virus; 1.
DR ProDom; PD001151; Porcine_RR_virus; 1.
SQ SEQUENCE 200 AA; 22427 MW; E8A3B9DA71AF88BB CRC64;

Query Match 76.6%; Score 36; DB 12; Length 200;
Best Local Similarity 77.8%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 STEGFFSGR 9
Db 97 STAGFFHGR 105

RESULT 4
O9J7K2 PRELIMINARY; PRT; 200 AA.
AC O9J7K2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Envelope glycoprotein.
OS Porcine reproductive and respiratory syndrome virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Arteriviridae; Arterivirus.
OX NCBI_TaxID=28344;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PRRSV23;
RX MEDLINE=20109045; PubMed=10640555;
RA Goldberg T.L.; Hahn E.C.; Weigel R.M.; Scherba G.;
RT "Genetic, geographical and temporal variation of porcine reproductive
RT and respiratory syndrome virus in Illinois.";
RL J. Gen. Virol. 81:171-179(2000).
DR EMBL; AF176445; AAF36251.1; -.
DR InterPro; IPR001332; Arteri_glycop.
DR InterPro; IPR003239; Porcine_RR_virus.
DR Pfam; PF00951; Arteri_glycop; 1.
DR ProDom; PD001151; Porcine_RR_virus; 1.
DR ProDom; PD001151; Porcine_RR_virus; 1.
SQ SEQUENCE 200 AA; 22379 MW; E47ED384FABF23AA CRC64;

Query Match 76.6%; Score 36; DB 12; Length 200;
Best Local Similarity 77.8%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 STEGFFSGR 9
Db 97 STAGFFHGR 105

RESULT 5
O990Q8 PRELIMINARY; PRT; 200 AA.
AC O990Q8;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE Envelope glycoprotein.
OS Porcine reproductive and respiratory syndrome virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Arteriviridae; Arterivirus.
OX NCBI_TaxID=28344;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=98-31701-1;
```


CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL; AE001268; AAC65957.1; -
 CC PIR; B71254; B71254.
 CC TIGR; TP1000; -
 CC KW Hypothetical protein; Complete proteome.
 CC SEQUENCE 223 AA; 25428 MW; 770AB73F904F530E CRC64;

Query Match 68.1%; Score 32; DB 1; Length 223;
 Best Local Similarity 66.7%; Pred. No. 26;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 STEGFFSGR 9
 |||||:|
 DB 158 STEGYQTGR 166

RESULT 14
 ACDA_BACSU STANDARD; PRT; 379 AA.
 AC P45867;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Acyl-CoA dehydrogenase (EC 1.3.99.-)
 GN ACDA OR ACD.
 OS Bacillus subtilis.
 OS Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID:1423;
 RN [1]_TaxID:1423;
 RP SEQUENCE FROM N.A.
 RC STRAIN:168;
 RX MEDLINE=98015417; PubMed=9353933;
 RA Prescan E., Moszer I., Boursier L., Cruz Ramos H.C., De La Fuente V.,
 RA Hullo M.-F., Lelong C., Schleich S., Sekowska A., Song B.H.,
 RA Villani G., Kunst F., Danchin A., Glaser P.
 RT "The Bacillus subtilis genome from gerBC (311 degrees) to lhcR (334
 RT degrees).";
 RL Microbiology 143:3313-3328(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN:168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Borries R., Boursier L., Brans A., Braun M., Brignall S.C., Bron S.,
 RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Conneron I.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
 RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Funa S., Galizzi A., Galleron N.,
 RA Chim S.Y., Glaser P., Goffeau A., Goughly E.J., Grandi G.,
 RA Guisepi G., Guy B.J., Haga K., Haeich J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetalle D., Porwollik S., Prescott A.M.,
 RA Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivalta C., Roche E., Roche B., Rose M., Sadate Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Solido B.,
 RA Sorokin A., Taconi E., Takagi T., Takahashi H., Takenaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 RA Tostato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,

RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 RT subtilis".
 RL Nature 390:249-256(1997)
 CC -1- CATALYTIC ACTIVITY: Acyl-CoA + ETF = 2,3-dehydroacyl-CoA + reduced
 CC ETF.
 CC -1- COFACTOR: FAD (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE ACYL-COA DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

EMBL; Z497B2; CAAG8968.1; -
 EMBL; Z99123; CAB15745.1; -
 PIR; S55421; S55421.
 HSSP; Q06319; 1BUC.
 DR Subtilisin; BG11239; acda.
 DR InterPro; IPR006089; Acyl-CoA dh.
 DR InterPro; IPR006090; Acyl-CoA dh C.
 DR InterPro; IPR006091; Acyl-CoA dh M.
 DR InterPro; IPR006092; Acyl-CoA dh N.
 DR Pfam; PF00441; Acyl-CoA dh; 1.
 DR Pfam; PF02770; Acyl-CoA dh; 1.
 DR Pfam; PF02771; Acyl-CoA dh; 1.
 DR PROSITE; PS00072; ACYL_COA_DH_1; 1.
 DR PROSITE; PS00073; ACYL_COA_DH_2; 1.
 KW Oxidoreductase; Flavoprotein; FAD; Complete proteome.
 SQ SEQUENCE 379 AA; 41446 MW; 4D09861D59718EF9 CRC64;

Query Match 68.1%; Score 32; DB 1; Length 379;
 Best Local Similarity 71.4%; Pred. No. 46;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 EGFPSGR 9
 |||||:
 DB 191 EGFFTCK 197

RESULT 15
 PPAL_PICPA STANDARD; PRT; 468 AA.
 AC P52291;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Acid phosphatase PHO1 precursor (EC 3.1.3.2).
 GN PHO1.
 OS Pichia pastoris (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Pichia.
 OX NCBI_TaxID:4922;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96001238; PubMed=7557473;
 RA Payne W.E., Gannon P.M., Kaiser C.A.;
 RT "An inducible acid phosphatase from the yeast Pichia pastoris:
 RT characterization of the gene and its product.";
 RL Gene 163:19-26(1995).
 CC -1- CATALYTIC ACTIVITY: An orthophosphoric monoester + H(2)O = an
 CC alcohol + phosphate.
 CC -1- INDUCTION: By phosphate starvation.
 CC -1- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

```

DR PHCT-2DPAGE; 084101;
DR InterPro; IPR000759; Adnrx_reductase.
DR InterPro; IPR001327; FAD_pyr_redux.
DR InterPro; IPR00103; Pyridine_redux.2.
DR InterPro; IPR005982; Thioresox_redux.
DR Pfam; PF00070; pyr_redux; 1.
DR PRINTS; PR00419; ADXRDTASE.
DR PRINTS; PR00368; PABNR.
DR PRINTS; PR00465; PNDRTASEII.
DR TIGRFAMS; TIGR01292; TRX_redux; 1.
DR PROSITE; PS00573; PYRIDINE_REDOX.2; 1.
KW Redox-active center; Oxidoreductase; NADP; Flavoprotein; FAD;
KW Complete proteome.
FT NP_BIND 4 19 FAD (ADP PART) (PROBABLE).
FT DISULFID 138 141 REDOX-ACTIVE (BY SIMILARITY).
FT NP_BIND 152 166 NAD(P) (BY SIMILARITY).
FT NP_BIND 273 283 FAD (FLAVIN PART) (BY SIMILARITY).
SQ SEQUENCE 312 AA; 33508 MW; 729D0D22F8FA0A39 CRC64;

Query Match 70.2%; Score 33; DB 1; Length 312;
Best Local Similarity 100.0%; Pred.No. 23;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EGFFSG 8
DB 33 EGFFSG 38

RESULT 10
METE_MYCTU
ID METE_MYCTU STANDARD; PRT; 759 AA.
AC O06584;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE 5-methyltetrahydropteroyltri-L-glutamate--homocysteine methyltransferase
DE (EC 2.1.1.14) (Methionine synthase, vitamin-B12 independent isozyme)
DE (Cobalamin-independent methionine synthase).
DE METE OR RV1133C OR W1185 OR MTC2268.22.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
-OX NCBI_TaxID=1773;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekai F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moulé S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CATALYZES THE TRANSFER OF A METHYL GROUP FROM 5-
CC METHYLTHETRAHYDROFOLATE TO HOMOCYSTEINE RESULTING IN METHIONINE
CC FORMATION (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: 5-methyltetrahydropteroyltri-L-glutamate + L-

```

```

CC homocysteine = tetrahydropteroyltri-L-glutamate + L-methionine.
CC -1- COPACTOR: ZINC; BINDS ONE ION PER SUBUNIT (BY SIMILARITY).
CC -1- PATHWAY: TERMINAL STEP IN THE DE NOVO BIOSYNTHESIS OF METHIONINE.
CC -1- SIMILARITY: Belongs to the vitamin-B12 independent methionine
CC synthase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z95585; CAB09044.1;
CC EMBL; A8006995; AAK45422.1;
CC PIR; F70539; F70539.
CC TIGR; W1185;
CC Tuberculin; Rv1133c;
CC HAMAP; MF_00172; 1.
CC InterPro; IPR006276; Met_syn_B12ind.
CC InterPro; IPR002629; Methionine synt.
CC Pfam; PF01717; Methionine synt; 1.
CC ProDom; PD004692; Methionine synt; 2.
CC TIGRFAMS; TIGR01371; met_syn_B12ind; 1.
KW Transferase; Methyltransferase; Methionine biosynthesis; Zinc; Repeat;
KW Complete proteome.
FT METAL 647 647 ZINC (BY SIMILARITY).
FT METAL 649 649 ZINC (BY SIMILARITY).
FT METAL 732 732 ZINC (BY SIMILARITY).
SQ SEQUENCE 759 AA; 81581 MW; 702F90BF79B25C8D CRC64;

Query Match 70.2%; Score 33; DB 1; Length 759;
Best Local Similarity 55.6%; Pred.No. 59;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 STEGPPSG 9
DB 29 ATEGYWAGR 37

RESULT 11
SYAC_SCHPO
ID SYAC_SCHPO STANDARD; PRT; 959 AA.
AC O13914;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable alanyl-tRNA synthetase, cytoplasmic (EC 6.1.1.7) (Alanine-
DE -tRNA ligase) (Alars).
GN SPAC23C11.09.
OS Schizosaccharomyces pombe (Fission Yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
-OX NCBI_TaxID=4896;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moulé S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford J., Rutter S., Saunders R., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grynoprez B.

```

RL J. Gen. Virol. 72:1505-1514(1991).
CC -1- SIMILARITY: IDENTICAL FOR THE FIRST 132 AA, AND 75.3% IDENTICAL
CC FOR THE NEXT 145 AA TO THE RNA1 POLYPROTEIN.
CC -1- SIMILARITY: TO THE RNA2 POLYPROTEIN OF OTHER NEPOVIRUSES.
CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-122 IS THE
CC INITIATOR.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; D12477; BAA02043.1; -
CC PIR; JQ1093; GNVVTR.
CC InterPro; IPR005054; Nepo_coat.
CC InterPro; IPR005305; Nepo_coat_C.
CC InterPro; IPR005306; Nepo_coat_N.
CC Pfam; PF03391; Nepo_coat_1.
CC Pfam; PF03688; Nepo_coat_C; 1.
CC Pfam; PF03689; Nepo_coat_N; 1.
CC Polyprotein; Coat protein; Repeat.
CC CHAIN 1321 1882 COAT PROTEIN (POTENTIAL).
FT DOMAIN 554 698 2.5 X TANDEM REPEATS, PRO-RICH.
FT REPEAT 554 606 1.
FT REPEAT 607 659 2.
FT REPEAT 660 698 3 (INCOMPLETE AND APPROXIMATE).
FT SEQUENCE 1882 AA; 206802 MW; 0F895B63AE8DD9D CRC64;
CC
CC Query Match 72.3%; Score 34; DB 1; Length 1882;
CC Best Local Similarity 66.7%; Pred. No. 95;
CC Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
CC
CC 1 STEGPFSCR 9
CC : |||:
CC 1717 SSGGFTGR 1725
CC
CC RESULT 6
CC VC02_SPKVA STANDARD; PRT; 92 AA.
CC ID VC02_SPKVA STANDARD; PRT; 92 AA.
CC AC F32230;
CC DT 01-OCT-1993 (Rel. 27, Created)
CC DT 01-OCT-1993 (Rel. 27, Last sequence update)
CC DT 01-NOV-1995 (Rel. 32, Last annotation update)
CC DE Hypothetical protein C2.
CC GN C2L OR K3R.
CC OS Swinepox virus (strain Kasza) (SPV).
CC OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
CC OC Suipoxvirus.
CC OX NCBI_TaxID=10277;
CC [1]
CC SEQUENCE FROM N.A.
CC MEDLINE=9406924; PubMed=8249275;
CC Massung R.F., Jayarama V., Moyer R.W.;
CC "DNA sequence analysis of conserved and unique regions of swinepox
CC virus: identification of genetic elements supporting phenotypic
CC observations including a novel G protein-coupled receptor
CC homologue."
CC RL Homology 197:511-528(1993).
CC -1- FUNCTION: HOMOLOG OF SHOPE FIBROMA VIRUS T4A ORF.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; L22013; AAC37869.1; -
CC

DR EMBL; L21931; AAC37874.1; -
KW Hypothetical protein.
SQ SEQUENCE 92 AA; 10799 MW; 5F08C066FA953379 CRC64;
Query Match 70.2%; Score 33; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 6.5;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 EGFFSG 8
DB 84 EGFFSG 89
RESULT 7
TRXB_CHLPN STANDARD; PRT; 311 AA.
ID TRXB_CHLPN STANDARD; PRT; 311 AA.
AC Q928M4; Q9CQJ8;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Thioresoxin reductase (EC 1.8.1.9) (TRXR).
GN TRXB OR CPN0314 OR CP0444.
OS Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CWL029;
RX MEDLINE=99206606; PubMed=10192388;
RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
"Comparative genomes of Chlamydia pneumoniae and C. trachomatis."
RT Nat. Genet. 21:385-389(1999).
RL [2]
RN SEQUENCE FROM N.A.
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Bruham R.C., Shen C., Gill S.R., Heidelberg J.F.,
White O., Hickey E.K., Peterson J., Uterback T., Berry K., Bass S.,
Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
Gwin M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
Eisen J., Fraser C.M.;
"Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
pneumoniae AR39."
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
"Comparison of whole genome sequences of Chlamydia pneumoniae J138
from Japan and CWL029 from USA."
RL Nucleic Acids Res. 28:2311-2314(2000).
CC -1- CATALYTIC ACTIVITY: Thioresoxin + NADP(+) = thioresoxin disulfide
CC + NADPH.
CC -1- COFACTOR: FAD.
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- MISCELLANEOUS: THE ACTIVE SITE IS A REDOX-ACTIVE DISULFIDE BOND.
CC -1- SIMILARITY: BELONGS TO THE PYRIDINE NUCLEOTIDE-DISULFIDE
CC OXIDOREDUCTASES CLASS-II.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AE001616; AAD18463.1; -
DR EMBL; AE002205; AAF38283.1; -

```

FT ZN_FING 29 70 RING-TYPE.
FT ZN_FING 102 143 B BOX-TYPE.
FT DOMAIN 181 250 COILED COIL (POTENTIAL).
FT DOMAIN 360 485 SPRY.
SQ SEQUENCE 486 AA; 56369 MW; 8EC3E584541P9A2 CRC64;

Query Match 76.6%; Score 36; DB 1; Length 488;
Best Local Similarity 77.8%; Pred. No. 9.1;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 STEGPFSGR 9
DQ 354 ATEGFTSGR 362

RESULT 2
ID RN23 HUMAN STANDARD; PRT; 518 AA.
AC Q9HCV9; Q9HCV9; B302.
DT 16-OCT-2001 (Rel. 40, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE RING finger protein 23 (Testis-abundant finger protein) (Tripartite
DE motif-containing protein 39).
GN TRIM39 OR RNF23 OR TRP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Testis;
RX MEDLINE=20462913; PubMed=11006080;
RA Orino A., Yamagishi T., Tomimaga N., Yamauchi Y., Hishinuma T.,
RA Okada K., Suzuki M., Sato M., Nogi Y., Suzuki H., Inoue S.,
RA Yoshimura K., Shimizu Y., Muramatsu M.;
RT "Molecular cloning of testis-abundant finger protein/ring finger
RT protein 23 (RNF23), a novel RING-B box-coiled coil-B30.2 protein on
RT the class I region of the human MHC.";
RL Biochem. Biophys. Res. Commun. 276:45-51(2000).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Colon;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh P.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Hilton E., Kettelman M., Vadan A., Rodriguez S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q9HCV9-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9HCV9-2; Sequence=VSP_005755;
CC Note=No experimental confirmation available;
CC -1- TISSUE SPECIFICITY: Ubiquitous; highly expressed in testis.
CC -1- SIMILARITY: Contains 1 RING-type zinc finger.

```

CC -1- SIMILARITY: Contains 1 B box-type zinc finger.
CC -1- SIMILARITY: Contains 1 SPRY domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC EMBL; AB046381; BAB16374.1; --
CC EMBL; BC007661; AAH07661.1; --
CC PIR; JC7387; JC7387.
CC HSP; P15919; IRMD.
CC Genew; HGNC:10065; TRIM39.
CC MIM; 605700; --
CC InterPro; IPR001870; B302.
CC InterPro; IPR006574; PRY.
CC InterPro; IPR003877; SPRY receptor.
CC InterPro; IPR000315; Znf_Ebox.
CC InterPro; IPR001841; Znf_Ring.
CC Pfam; PF00622; SPRY; 1.
CC Pfam; PF00643; zf-B_box; 1.
CC Pfam; PF00097; zf-C3HC4; 1.
CC SMART; SM00336; BBOX; 1.
CC SMART; SM00589; PRY; 1.
CC SMART; SM00184; RING; 1.
CC SMART; SM00449; SPRY; 1.
CC PROSITE; PS01119; ZF_BBOX; 1.
CC PROSITE; PS00518; ZF_RING_1; 1.
CC PROSITE; PS0089; ZF_RING_2; 1.
CC Zinc-finger; Coiled coil; Alternative splicing.
CC ZN_FING 29 70 RING-TYPE.
CC ZN_FING 102 143 B BOX-TYPE.
CC DOMAIN 181 250 COILED COIL (POTENTIAL).
CC DOMAIN 360 515 SPRY.
CC VARSPLIC 269 298 Missing (in isoform 2).
CC FTID=VSP_005755.
CC CONFLICT 137 137 P -> A (IN REF. 2).
CC SEQUENCE 518 AA; 59716 MW; DA92B328F253B828 CRC64;
Query Match 76.6%; Score 36; DB 1; Length 518;
Best Local Similarity 77.8%; Pred. No. 9.7;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 STEGPFSGR 9
DQ 384 ATEGFTSGR 392
RESULT 3
ID UBPY_CAEEL STANDARD; PRT; 1302 AA.
AC Q09931;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Probable ubiquitin carboxyl-terminal hydrolase K02C4.3 (EC 3.1.2.15)
DE (Ubiquitin thiolesterase) (Ubiquitin-specific processing protease)
DE (Deubiquitinating enzyme).
GN K02C4.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
CX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Lightening J.;
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Ubiquitin C-terminal thiolester + H(2)O =
CC ubiquitin + a thiol.

RESULT 12

H64825
hypothetical protein b0872 - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
C:Accession: H64825
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Cohen, D.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: H64825
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-322 <BLAT>
A:Cross-references: GB:AE000189; GB:U00096; NID:gl787097; PIDN:AACT73959.1; PID:gl787098;
A:Experimental source: strain K-12, substrain MG1655
C:Superfamily: phthalate dioxygenase reductase; cytochrome-b5 reductase homology; ferredoxin
P:258-312/Domain: ferredoxin [2Fe-2S] homology <FER>

Query Match 72.3%; Score 34; DB 2; Length 322;
Best Local Similarity 75.0%; Pred. No. 34;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 TEGFFSGR 9
|||||
Db 174 TEGFTAGR 181

RESULT 13

T26902
hypothetical protein Y44F5A.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T26902
R:McMurray, A.
submitted to the EMBL Data Library, January 1998
A:Reference number: Z20284
A:Accession: T26902
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-383 <WIL>
A:Cross-references: EMBL:AL021180; PIDN:CAAL5980.1; GSPDB:GN00021; CESP:Y44F5A.1
A:Experimental source: clone Y44F5A
C:Genetics:
A:Gene: CESP:Y44F5A.1
A:Map position: 3
A:Introns: 85/2; 179/3; 283/3

Query Match 72.3%; Score 34; DB 2; Length 383;
Best Local Similarity 85.7%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TEGFFSG 8
|||||
Db 329 TEGFFDG 335

RESULT 14

GNVTR
genome polyprotein 2 - tomato ringspot virus (strain raspberry)
N:Contains: coat protein
C:Species: tomato ringspot virus
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jun-2000
C:Accession: JQ1093
R:Rott, M.E.; Tremaine, J.H.; Roehon, D.M.
J. Gen. Virol. 72, 1505-1514, 1991
A:Title: Nucleotide sequence of tomato ringspot virus RNA-2.
A:Reference number: JQ1093; MUID:91311402; PMID:1856689
A:Accession: JQ1093
A:Molecule type: genomic RNA
A:Residues: 1-1882 <RGT>

A:Cross-references: GB:D12477; GB:D01129; NID:G222674; PIDN:BAA02043.1; PID:G222675
A>Note: It is uncertain whether Met-1 or Met-122 is the initiator
C:Genetics:

A:Map position: segment 2
C:Superfamily: tomato ringspot virus genome polyprotein
C:KeyWords: coat protein; glycoprotein; polyprotein
F:1321-1882/Product: coat protein #status predicted <MAT>
F:269,295,1183,1316,1543,1561,1735/Binding site: carbohydrate (Asn) (covalent) #status

Query Match 72.3%; Score 34; DB 1; Length 1882;
Best Local Similarity 66.7%; Pred. No. 2e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 STEGPFSGR 9
|:|||||
Db 1717 SSGTFTGR 1725

RESULT 15

B86530
thioredoxin reductase [imported] - Chlamydomonas reinhardtii (strain J138)
C:Species: Chlamydomonas reinhardtii, Chlamydia pneumoniae
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C:Accession: B86530
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; I
Nucleic Acids Res. 28, 2311-2314, 2000
A:Title: Comparison of whole genome sequences of Chlamydia pneumoniae J138.
A:Reference number: A86491; MUID:20330349; PMID:10871362
A:Accession: B86530
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-311 <STO>
A:Cross-references: GB:BA000008; NID:98978688; PIDN:BAA98524.1; GSPDB:GN00142
A:Experimental source: strain J138
C:Genetics:
A:Gene: trxB
C:Superfamily: thioredoxin reductase; thioredoxin reductase homology

Query Match 70.2%; Score 33; DB 2; Length 311;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EGFFSG 8
|||||
Db 33 EGFFSG 38

Search completed: December 11, 2003, 18:29:31
Job time : 22.3333 secs

A/Cross-references: GB:AE002239; GB:AE002161; NID:g7189708; PIDN:AAF38592.1; PID:g718970
 A/Experimental source: strain AR39, HL cells
 C/Genetics:
 A/Gene: CPN1057; CP0793
 C/Superfamily: conserved hypothetical protein yya1

Query Match 80.9%; Score 38; DB 2; Length 700;
 Best Local Similarity 77.8%; Pred. No. 11;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 STEGFFSQR 9
 | | | | |
 Db 367 SREGFFNGR 375

RESULT 3

D83555
 A/Title: Hypothetical protein PA0716 [imported] - Pseudomonas aeruginosa (strain PA01)
 C/Species: Pseudomonas aeruginosa
 C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C/Accession: D83555
 R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
 ; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
 A/Reference number: A82950; MUID:20437337; PMID:10984043
 A/Accession: D83555
 A/Status: Preliminary
 A/Molecule type: DNA
 A/Residues: 1-441 <STO>
 A/Cross-references: GB:AE004507; GB:AE004091; NID:g9946596; PIDN:AAG04105.1; GSPDB:GN001
 A/Experimental source: strain PA01
 C/Genetics:
 A/Gene: PA0716

Query Match 76.6%; Score 36; DB 2; Length 441;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STEGFFS 7
 | | | | |
 Db 335 STEGFFS 341

RESULT 4

JC7387
 A/Title: testis-abundant finger protein - human
 C/Species: Homo sapiens (man)
 C/Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 08-Dec-2000
 C/Accession: JC7387
 R/Otomo, A.; Yamagishi, T.; Tominga, N.; Yamauchi, Y.; Hishinuma, T.; Okada, K.; Suzuki
 Biochem. Biophys. Res. Commun. 276, 45-51, 2000
 A/Title: Molecular cloning of testis-abundant finger protein/ring finger protein 23 (RNF
 A/Reference number: JC7387
 A/Contents: Testis
 A/Accession: JC7387
 A/Molecule type: mRNA
 A/Residues: 1-518 <ORI>
 A/Cross-references: DDBJ:AB046381
 C/Comment: This protein, a member of the ring-B box-coiled coil-B30.2 protein family, p
 C/Genetics:
 A/Gene: ttfp
 A/Keywords: coiled coil; testis

Query Match 76.6%; Score 36; DB 2; Length 518;
 Best Local Similarity 77.8%; Pred. No. 22;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 STEGFFSQR 9

Db 384 ATEGFTSGR 392
 : | | | | |

RESULT 5

AD3057
 A/Title: glycoen debranching enzyme glgX [imported] - Agrobacterium tumefaciens (strain C58, D
 C/Species: Agrobacterium tumefaciens
 C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
 C/Accession: AD3057
 R/Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo,
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClel
 ; Karp, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
 ster, R.W.
 A/Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
 A/Reference number: AB2577; MUID:21608550; PMID:11743193
 A/Accession: AD3057
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-651 <KUR>
 A/Cross-references: GB:AE008689; PIDN:AAL44874.1; PID:g17742522; GSPDB:GN00187
 A/Experimental source: strain C58 (Dupont)
 C/Genetics:
 A/Gene: glgX
 A/Map position: linear chromosome
 C/Superfamily: glyX protein

Query Match 76.6%; Score 36; DB 2; Length 651;
 Best Local Similarity 77.8%; Pred. No. 27;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 STEGFFSQR 9
 | | | | |
 Db 556 SESGFFSQR 564

RESULT 6

B98229
 A/Title: glycoen debranching enzyme (AJ291603) [imported] - Agrobacterium tumefaciens (strain C
 C/Species: Agrobacterium tumefaciens
 C/Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 18-Nov-2002
 C/Accession: B98229
 R/Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldmar
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.
 Science 294, 2323-2328, 2001
 A/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium t
 A/Reference number: A97359; MUID:21608551; PMID:11743194
 A/Accession: B98229
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-651 <KUR>
 A/Cross-references: GB:AE007870; PIDN:AAK99356.1; PID:g15159204; GSPDB:GN00170
 C/Genetics:
 A/Gene: AGR L 1566
 A/Map position: linear chromosome
 C/Superfamily: glyX protein

Query Match 76.6%; Score 36; DB 2; Length 651;
 Best Local Similarity 77.8%; Pred. No. 27;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 STEGFFSQR 9
 | | | | |
 Db 556 SESGFFSQR 564

RESULT 7

AH0508
 A/Title: NADH oxidoreductase Hcr (SC 1.-.-.-) [imported] - Salmonella enterica subsp. enterica
 C/Species: Salmonella enterica subsp. enterica serovar Typhi
 A/Note: this species has also been called Salmonella typhi

STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION NUMBER: US/09/828,326
FILING DATE: 05-Apr-2001

CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/341,555
FILING DATE: <Unknown>
APPLICATION NUMBER: US 07/891,177
FILING DATE: 29-MAY-1992

ATTORNEY/AGENT INFORMATION:
NAME: Storella, John R.
REGISTRATION NUMBER: 32,944
REFERENCE/DOCKET NUMBER: 016866-000200US

TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear

MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-828-326-20

Query Match 68.1%; Score 32; DB 11; Length 14;
Best Local Similarity 85.7%; Pred. No. 9.3;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 EGFPSGR 9
DB 8 EGFPSAR 14

RESULT 10
US-10-142-935-11
; Sequence 11, Application US/10142935
; Publication No. US2003004418A1
; GENERAL INFORMATION:
; APPLICANT: DAVIS, Stacey
; TITLE OF INVENTION: METHOD AND COMPOSITIONS FOR INHIBITING THROMBIN-INDUCED COAGULATION
; FILE REFERENCE: P07201US01/BAS
; CURRENT APPLICATION NUMBER: US/10/142,935
; CURRENT FILING DATE: 2002-05-13
; PRIOR APPLICATION NUMBER: US 60/290,072
; PRIOR FILING DATE: 2001-05-13
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patent in version 3.1

SEQ ID NO 11
LENGTH: 14
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-10-142-935-11

Query Match 68.1%; Score 32; DB 15; Length 14;
Best Local Similarity 85.7%; Pred. No. 9.3;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 EGFPSGR 9
DB 8 EGFPSAR 14

RESULT 11
US-10-142-935-8
; Sequence 8, Application US/10142935
; Publication No. US2003004418A1
; GENERAL INFORMATION:
; APPLICANT: DAVIS, Stacey
; TITLE OF INVENTION: METHOD AND COMPOSITIONS FOR INHIBITING THROMBIN-INDUCED COAGULATION
; FILE REFERENCE: P07201US01/BAS
; CURRENT APPLICATION NUMBER: US/10/142,935
; CURRENT FILING DATE: 2002-05-13
; PRIOR APPLICATION NUMBER: US 60/290,072
; PRIOR FILING DATE: 2001-05-13
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patent in version 3.1

SEQ ID NO 8
LENGTH: 15
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-10-142-935-8

Query Match 68.1%; Score 32; DB 15; Length 15;
Best Local Similarity 85.7%; Pred. No. 10;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 EGFPSGR 9
DB 3 EGFPSAR 9

RESULT 12
US-10-197-954-54
; Sequence 54, Application US/10197954
; Publication No. US20030119021A1
; GENERAL INFORMATION:
; APPLICANT: K'ster, Hubert
; APPLICANT: Siddiqi, Suhaib
; APPLICANT: Little, Daniel
; TITLE OF INVENTION: Capture Compounds, Collections Thereof
; TITLE OF INVENTION: And Methods For Analyzing The Proteome And Complex
; TITLE OF INVENTION: Compositions
; FILE REFERENCE: 24743-2305
; CURRENT APPLICATION NUMBER: US/10/197,954
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: 60/306,019
; PRIOR FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 60/314,123
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: 60/363,433
; PRIOR FILING DATE: 2002-03-11
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 54
LENGTH: 15
TYPE: PRT
ORGANISM: Homo Sapien
US-10-197-954-54

Query Match 68.1%; Score 32; DB 15; Length 15;
Best Local Similarity 85.7%; Pred. No. 10;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 EGFPSGR 9
DB 9 EGFPSAR 15

RESULT 13
US-10-142-935-5
; Sequence 5, Application US/10142935
; Publication No. US2003004418A1

; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1405
; LENGTH: 127
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1405

Query Match 74.5%; Score 35; DB 9; Length 127;
Best Local Similarity 85.7%; Pred. No. 25;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 EGFYSGR 9
|||:|
Db 2 EGFYSGR 8
|||:|

RESULT 3
US-09-738-626-5763
; Sequence 5763, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENO, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OKAZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738, 626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 5763
; LENGTH: 321
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-5763

Query Match 72.3%; Score 34; DB 10; Length 321;
Best Local Similarity 85.7%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 EGFYSGR 9
|||:|
Db 275 EGFYSGR 281
|||:|

RESULT 4
US-10-237-386-66
; Sequence 66, Application US/10237386
; Publication No. US20030180895A1
; GENERAL INFORMATION:
; APPLICANT: Danisco A/S
; APPLICANT: Sibiessen, Ole
; APPLICANT: Sorensen, Jens
; TITLE OF INVENTION: Xylanase Variants Having Altered Sensitivity to Xylanase Inhibitor
; FILE REFERENCE: 674509-2046

; CURRENT APPLICATION NUMBER: US/10/237,386
; CURRENT FILING DATE: 2002-12-06
; PRIOR APPLICATION NUMBER: PCT/IB01/00426
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: GB 0005585.5
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: GB 0015751.1
; PRIOR FILING DATE: 2000-06-27
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 66
; LENGTH: 280
; TYPE: PRT
; ORGANISM: P. stipitis
US-10-237-386-66

Query Match 70.2%; Score 33; DB 12; Length 280;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 STEGFFS 7
|||:|
Db 240 STEGFFS 246
|||:|

RESULT 5
US-10-032-201B-203
; Sequence 203, Application US/10032201B
; Publication No. US20030167524A1
; GENERAL INFORMATION:
; APPLICANT: Van Rooijen, Gijb
; APPLICANT: Deckers, Harm
; APPLICANT: Heifetz, Peter Bernard
; APPLICANT: Briggs, Steven
; APPLICANT: Dalmla, Bipin Kumar
; APPLICANT: Del Val, Greg
; APPLICANT: Moloney, Maurice
; APPLICANT: Zaplachinski, Steve
; TITLE OF INVENTION: METHODS FOR THE PRODUCTION OF MULTIMERIC PROTEINS, AND RELATED
; FILE REFERENCE: 38814 351B
; CURRENT APPLICATION NUMBER: US/10/032,201B
; CURRENT FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 313
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 203
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-10-032-201B-203

Query Match 70.2%; Score 33; DB 12; Length 311;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EGFPSG 8
|||:|
Db 33 EGFPSG 38
|||:|

RESULT 6
US-10-032-201B-202
; Sequence 202, Application US/10032201B
; Publication No. US20030167524A1
; GENERAL INFORMATION:
; APPLICANT: Van Rooijen, Gijb
; APPLICANT: Deckers, Harm
; APPLICANT: Heifetz, Peter Bernard
; APPLICANT: Briggs, Steven
; APPLICANT: Dalmla, Bipin Kumar
; APPLICANT: Del Val, Greg
; APPLICANT: Zaplachinski, Steve
; APPLICANT: Moloney, Maurice


```

; FILING DATE: 25-JUL-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Baron, Ronald J
; REGISTRATION NUMBER: 29281
; REFERENCE/DOCKET NUMBER: 454-15
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516 822 3550
; TELEFAX: 516 822 3582
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-900-660A-4

```

```

Query Match      68.1%; Score 32; DB 2; Length 14;
Best Local Similarity 85.7%; Pred. No. 3.7;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 3 EGFPSGR 9
DB 8 EGFPSAR 14

```

RESULT 15

```

US-08-900-660A-5
; Sequence 5, Application US/08900660A
; Patent No. 5876947

```

```

; GENERAL INFORMATION:
; APPLICANT: Kudryk, Bohdan J
; APPLICANT: Bini, Alessandra
; APPLICANT: Zhang, Jian-Zhong
; TITLE OF INVENTION: MONOSPECIFIC ANTIBODY REACTIVE
; TITLE OF INVENTION: WITH FIBRINOGEN AND
; TITLE OF INVENTION: FIBRINOPEPTIDE B
; NUMBER OF SEQUENCES: 6

```

```

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: USA
; ZIP: 11753

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/900,660A

```

```

; FILING DATE: 25-JUL-1997
; CLASSIFICATION: 435

```

```

; ATTORNEY/AGENT INFORMATION:
; NAME: Baron, Ronald J

```

```

; REGISTRATION NUMBER: 29281
; REFERENCE/DOCKET NUMBER: 454-15
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516 822 3550

```

```

; TELEFAX: 516 822 3582
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids

```

```

; TYPE: amino acid
; TOPOLOGY: linear

```

```

; MOLECULE TYPE: peptide
; FEATURE:

```

```

; NAME/KEY: modified site
; LOCATION: 1

```

```

; OTHER INFORMATION: /note="pyroglutamic acid"
US-08-900-660A-5

```

```

Query Match      68.1%; Score 32; DB 2; Length 14;
Best Local Similarity 85.7%; Pred. No. 3.7;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 3 EGFPSGR 9
DB 8 EGFPSAR 14

```

```

Search completed: December 11, 2003, 18:30:34
Job time : 10.2222 secs

```

US-09-328-352-5503

Query Match
Best Local Similarity 70.2%; Score 33; DB 4; Length 3892;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EGFSSG 8
DB 69 EGFSSG 74

RESULT 10
US-08-592-500-41
; Sequence 41, Application US/08592500
; Patent No. 6005089
; GENERAL INFORMATION:
; APPLICANT: Lanza, Francois
; APPLICANT: Phillips, David R.
; APPLICANT: Cazenave, Jean-Pierre
; TITLE OF INVENTION: Platelet Glycoprotein V Gene and Uses
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Khourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 09-JUL-1993
; APPLICATION NUMBER: US/08/592,500
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/089,455
; FILING DATE: 09-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 12418-28
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..11
; OTHER INFORMATION: /note= "Amino acid sequence of the
; OTHER INFORMATION: human fibrinogen (Fg) B-beta chain thrombin
; OTHER INFORMATION: cleavage site."
; FEATURE:
; NAME/KEY: Region
; LOCATION: 7..8
; OTHER INFORMATION: /note= "Amino acid residues
; OTHER INFORMATION: identical to GPV."
; FEATURE:
; NAME/KEY: Region
; LOCATION: 11
; OTHER INFORMATION: /note= "Amino acid residue
; OTHER INFORMATION: identical to GPV."
US-08-592-500-41

Query Match 68.1%; Score 32; DB 3; Length 11;

Best Local Similarity 85.7%; Pred. No. 2.9;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 EGFSSG 9
DB 1 EGFSSG 7

RESULT 11
US-08-195-006-41
; Sequence 41, Application US/08195006
; Patent No. 6083688
; GENERAL INFORMATION:
; APPLICANT: Lanza, Francois
; APPLICANT: Phillips, David R.
; APPLICANT: Cazenave, Jean-Pierre
; TITLE OF INVENTION: Platelet Glycoprotein V Gene and Uses
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Khourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/195,006
; FILING DATE: 10-FEB-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/089,455
; FILING DATE: 09-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 12418-28
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..11
; OTHER INFORMATION: /note= "Amino acid sequence of the
; OTHER INFORMATION: human fibrinogen (Fg) B-beta chain thrombin
; OTHER INFORMATION: cleavage site."
; FEATURE:
; NAME/KEY: Region
; LOCATION: 7..8
; OTHER INFORMATION: /note= "Amino acid residues
; OTHER INFORMATION: identical to GPV."
; FEATURE:
; NAME/KEY: Region
; LOCATION: 11
; OTHER INFORMATION: /note= "Amino acid residue
; OTHER INFORMATION: identical to GPV."
US-08-195-006-41

Query Match 68.1%; Score 32; DB 3; Length 11;
Best Local Similarity 85.7%; Pred. No. 2.9;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

PRIOR APPLICATION NUMBER: 60/056,877	PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,889	PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,893	PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,930	PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,878	PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,662	PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,872	PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,882	PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,637	PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,903	PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,888	PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,879	PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,880	PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,894	PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,911	PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,636	PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,874	PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,910	PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,864	PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,631	PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,845	PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,892	PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/057,761	PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/047,595	PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,599	PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,588	PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,585	PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,586	PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,590	PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,594	PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,589	PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,593	PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,614	PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/043,578	PRIOR FILING DATE: 1997-04-11
PRIOR APPLICATION NUMBER: 60/043,576	PRIOR FILING DATE: 1997-04-11
PRIOR APPLICATION NUMBER: 60/047,501	PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/043,670	PRIOR FILING DATE: 1997-05-23

```
; Sequence 30, Application US/09970711
; Patent No. US20020081279A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Adam
; APPLICANT: Cotten, Matthew
; APPLICANT: Chiocca, Susanna
; APPLICANT: Kurzbaumer, Robert
; APPLICANT: Schaffner, Gotthold
; TITLE OF INVENTION: Chicken Embryo Lethal Orphan (CELO) Virus
; FILE REFERENCE: 0692.1800001
; CURRENT APPLICATION NUMBER: US/09/970,711
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/171,461
; PRIOR FILING DATE: 1998-01-12
; PRIOR APPLICATION NUMBER: PCT/EP97/01944
; PRIOR FILING DATE: 1997-04-18
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 30
; LENGTH: 575
; TYPE: PRT
; ORGANISM: CELO VIRUS
; FEATURE:
; OTHER INFORMATION: Position: 10269..11996/Product: E2b pTP
; US-09-970-711-30

Query Match      78.0%; Score 32; DB 9; Length 575;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 TALEAIG 8
DB      354 TALEAIG 360

RESULT 3
US-10-128-714-8018
; Sequence 8018, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wenqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroskin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 8018
; LENGTH: 974
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-10-128-714-8018

Query Match      78.0%; Score 32; DB 15; Length 974;
Best Local Similarity 77.8%; Pred. No. 3e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 STALEAIG 9
DB      24 STALDASGR 32

RESULT 4
US-09-880-505-135
; Sequence 135, Application US/09880505
; Publication No. US20030007976A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Tan, Paul L. J.
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: Methods and Compounds for the Treatment
; FILE REFERENCE: 11000.1007C2
; CURRENT APPLICATION NUMBER: US/09/880,505
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 09/324,542
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: US 08/997,080
; PRIOR FILING DATE: 1997-12-23
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 135
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Mycobacterium vaccae
US-09-880-505-135

Query Match      75.6%; Score 31; DB 11; Length 247;
Best Local Similarity 87.5%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 STALEAIG 8
DB      46 STATEAIG 53

RESULT 5
US-10-051-643-135
; Sequence 135, Application US/10051643
; Publication No. US20020197265A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Tan, Paul L. J.
; TITLE OF INVENTION: Methods and Compounds for the Treatment
; FILE REFERENCE: 11000.1008C2
; CURRENT APPLICATION NUMBER: US/10/051,643
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US09/156,181
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: US 08/996,624
; PRIOR FILING DATE: 1997-12-23
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 135
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Mycobacterium vaccae
US-10-051-643-135

Query Match      75.6%; Score 31; DB 14; Length 247;
Best Local Similarity 87.5%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 STALEAIG 8
DB      46 STATEAIG 53
```

XX 18-NOV-1999.
 PD
 XX
 PF 12-MAY-1999; 99WO-US10549.
 XX
 PR 12-MAY-1998; 98US-0085197.
 PR 05-MAY-1999; 99US-0085197.
 XX
 XX (BOCK/) BOCK S C.
 PA (PICA/) PICARD V.
 PA (ZEND/) ZENDEHROUH P.
 XX
 XX Bock SC, Picard V, Zendeherouh P;
 XX WPI; 2000-116274/10.
 DR
 XX
 XX New modified human antithrombin III compounds, used for treating e.g.
 PT sepsis, trauma, acute respiratory distress syndrome, restenosis,
 PT thrombosis, thromboembolism or stroke -
 XX
 PS Claim 13; Page 57; 75pp; English.
 XX
 CC The present sequence is from an antithrombin III (ATIII) variant, Bb
 CC derived from human ATIII.N135A cDNA insert of the pBlueBac baculovirus
 CC expression construct and comprises residues 385-393. The variant has
 CC improved resistance to elastase and Irg-activated neutrophils while
 CC retains anti-thrombin and anti-factor Xa activities. It may be
 CC expressed as glycoforms with enhanced heparin affinity which target the
 CC blood vessel wall more efficiently than ATIIIs with normal heparin
 CC affinity. The modified ATIIIs can be used to treat thrombin activation-
 CC related pathological symptoms due to sepsis, trauma, acute
 CC respiratory distress syndrome, restenosis, thrombosis, thromboembolism
 CC and stroke. It can also be used to reduce the risk of reocclusion
 CC and restenosis in percutaneous transluminal coronary angioplasty,
 CC thrombosis associated with surgery, ischaemia/reperfusion injury, and
 CC coagulation abnormalities in cancer or surgical patients.
 XX
 SQ Sequence 9 AA;
 Query Match 100.0%; Score 47; DB 21; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e-05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 STEGFFSGR 9
 Db | | | | | | | |
 1 STEGFFSGR 9
 RESULT 2
 AAY35714
 ID AAY35714 standard; Protein; 679 AA.
 XX
 AC AAY35714;
 XX
 DT 13-SEP-1999 (first entry)
 XX
 DE Amino acid sequence of a Chlamydia pneumoniae protein.
 KW
 KW Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
 KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
 KW vaccine; neutralising epitope.
 XX
 OS Chlamydia pneumoniae.
 XX
 XX WO9927105-A2.
 FN
 PD 03-JUN-1999.
 XX
 PF 20-NOV-1998; 98WO-IB01890.
 XX
 XX 04-NOV-1998; 96US-0107078.
 PR 21-NOV-1997; 97FR-0014673.
 PR
 XX

PA (GEST) GENSET.
 XX
 PI Griffais R;
 XX
 DR WPI, 1999-357842/30.
 XX
 FT Genome sequence of Chlamydia pneumoniae
 XX
 PS Page 1418-1419; Disclosure; 1912pp; English.
 XX
 CC AAY34584-Y35879 represent the proteins encoded by all the open reading
 CC frames in the complete genome (see AAY31990) of Chlamydia pneumoniae.
 CC C. pneumoniae causes respiratory disease such as pneumonia and
 CC bronchitis and is thought to be a contributing factor in heart
 CC disease, sarcoidosis, sinusitis, purulent otitis media, erythema
 CC nodosum or pharyngitis. The polypeptides encoded by the open reading
 CC frames of the C. pneumoniae genome (see AAY34584-Y35879) can be used in
 CC immunogenic compositions as vaccines. Vectors containing C. pneumoniae
 CC nucleotide sequences can also be used as immunogenic compositions,
 CC especially where the vector directs the expression of a neutralising
 CC epitope of C. pneumoniae.
 XX
 SQ Sequence 679 AA;
 Query Match 80.9%; Score 38; DB 20; Length 679;
 Best Local Similarity 77.8%; Pred. No. 58;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 STEGFFSGR 9
 Db | | | | | | | |
 367 SREGFFNGR 375
 RESULT 3
 AAU17466
 ID AAU17466 standard; Protein; 194 AA.
 XX
 AC AAU17466;
 XX
 DT 07-NOV-2001 (first entry)
 XX
 DE Novel signal transduction pathway protein, Seq ID 1031.
 XX
 KW Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;
 KW antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;
 KW immune system disorder; rheumatoid arthritis; inflammatory condition;
 KW organ transplant rejection; infection; hepatitis C; blood disorder;
 KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
 KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
 KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
 KW reproductive system; gastrointestinal; liver disorder; AIDS;
 KW acquired immune deficiency syndrome.
 XX
 OS Homo sapiens.
 XX
 XX WO200154733-A1.
 FN
 PD 02-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US01312.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.

CC other blood-related disorders (sickle cell anaemia), myeloproliferative
 CC disorders, primary haematopoietic disorders, hyperproliferative
 CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative
 CC disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal
 CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal
 CC disorders (e.g. glomerulonephritis), cardiovascular disorders
 CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in
 CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.
 CC Addison's disease), reproductive system disorders, gastrointestinal
 CC disorder (inflammatory disorders), liver disorders (cirrhosis).
 CC as stimulators of B-cell responsiveness to pathogens, activators of
 CC T-cells, to induce higher affinity antibodies, and as a means to induce
 CC tumour proliferation in pathologies e.g. acquired immune deficiency
 CC syndrome (AIDS). AAU17059-AAU17683 represent novel signal transduction
 CC pathway protein, amino acid sequences of the invention.
 XX

Query Match 76.6%; Score 36; DB 22; Length 194;
 Best Local Similarity 77.8%; Pred. No. 39;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 STEGFPSGR 9
 Db 60 ATEGTSGR 68

RESULT 4

AA58673
 ID AAY58673 standard; Protein; 200 AA.

AC AAY58673;

DT 25-APR-2000 (first entry)

XX Porcine reproductive respiratory syndrome virus protein.

DE PRRS; racoonpox virus; RPV; vaccine.

XX Porcine reproductive respiratory syndrome virus.

OS WO200003030-A2.

PN 20-JAN-2000.

XX 09-JUL-1999; 99WO-US15565.

XX 10-JUL-1998; 98US-0113750.

XX (SCHE) SCHERING-PLOUGH LTD.

FA Cochran MD, Junker DS;

XX WPI; 2000-171150/15.

XX N-PSDB; AAZ58057.

CC New recombinant racoonpox virus containing foreign DNA inserted into a
 CC non-essential region within the HindIII U genomic region, useful as a
 CC vaccine against pathogens in mammalian and avian species

XX Disclosure; Page 151-152; 164pp; English.

XX The present sequence is that of a protein encoded by an open
 CC reading frame (ORF) in genomic DNA (see AAZ58057) of porcine
 CC reproductive respiratory syndrome (PRRS) virus Eichelberger strain.
 CC The invention provides recombinant racoonpox virus (RPV)
 CC containing foreign DNA inserted into a non-essential region within
 CC its genomic DNA. In a particular embodiment, the foreign DNA may be
 CC that of PRRS ORF2, ORF3, ORF4, ORF5, ORF6 or ORF7. Homology vectors
 CC of the invention have a marker gene and a PRRS ORF flanked by RPV
 CC DNA. The vectors are constructed for the purpose of inserting
 CC foreign DNA into RPV. The recombinant RPVs are used in vaccines to
 CC protect against disease.

SQ Sequence 200 AA;

Query Match 76.6%; Score 36; DB 21; Length 200;
 Best Local Similarity 77.8%; Pred. No. 40;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 STEGFPSGR 9
 Db 97 STAGFFHGR 105

RESULT 5

AA43960

ID AAB43960 standard; Protein; 127 AA.

XX AAB43960;

DT 08-FEB-2001 (first entry)

DE Human cancer associated protein sequence SEQ ID NO:1405.

XX Human; cancer associated gene; cancer antigen; detection; cancer;
 KW diagnosis; cytostatic; proliferative; vulnery; immunomodulator;
 KW antidiabetic; antiasthmatic; antirheumatic; antidiabetic; antiviral;
 KW antiinflammatory; antithyroid; antiallergic; antibacterial; cardiac;
 KW dermatological; neuroprotective; thrombolytic; coagulant; neotropic;
 KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
 KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
 KW allergic reaction; graft versus host disease; organ rejection;
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;
 KW neurological disease; drug screening.

XX Homo sapiens.

XX WO200005350-A1.

XX 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US05882.

XX 12-MAR-1999; 99US-0124270.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;

XX WPI; 2000-587533/55.

XX N-PSDB; AAC78169.

XX Novel isolated nucleic acids comprising sequences encoding peptides
 XX useful for treating or diagnosing e.g. cancer -

XX Claim 11; Page 2083-2084; 2352pp; English.

XX AAC77607 to AAC78448 encode the human cancer associated proteins given
 CC in AA43398 to AAB44239. The proteins can have activities based on the
 CC tissues and cells the genes are expressed in. Example of activities
 CC include: cytostatic; proliferative; vulnery; immunomodulator;
 CC antidiabetic; antiasthmatic; antirheumatic; antidiabetic; antiviral;
 CC antiinflammatory; antithyroid; antiallergic; antibacterial; coagulant;
 CC dermatological; neuroprotective; cardiac; thrombolytic; antidiabetic;
 CC vasotropic; antipsoriatic; antipsoriatic and antiangiogenic. The
 CC polynucleotides and polypeptides can be used for preventing, treating or
 CC ameliorating medical conditions and diagnosing pathological conditions.
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
 CC the present invention may be used to treat immune disorders by activating
 CC or inhibiting the proliferation, differentiation or mobilisation of
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune
 CC disorders, allergic reactions, graft versus host disease, and organ
 CC rejection, modulate haemostatic or thrombolytic activity, modulate
 CC inflammation, cancers, cardiovascular disorders, neurological disease and
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to

```

Query Match      72.3%; Score 34; DS 22; Length 117;
Best Local Similarity 75.0%; Pred. NO. 57;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY      2 TEGFFSGR 9
        ||||| :||
Db      61 TEGFIAGR 68

```

QY 2 TEGFESGR 9
||| : ||
Db 153 TEGFIAGR 160

153 TEGFIAGR 160

153 TEGFIAGR 160

KW	Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;	PR	14-SEP-2000;	2000US-0233064.
KW	ligase; hyperproliferative disorder; immunodeficiency disorder;	PR	14-SEP-2000;	2000US-0233065.
KW	autoimmune disorder; neurological disorder; metabolic disorder;	PR	21-SEP-2000;	2000US-0234223.
KW	inflammatory disorder; cardiovascular disorder; reproductive disorder;	PR	21-SEP-2000;	2000US-0234274.
KW	blood-related disorder; infectious disorder; cytostatic; anti arthritic;	PR	23-SEP-2000;	2000US-0234997.
KW	nephrotropic; anticoagulant.	PR	25-SEP-2000;	2000US-0234998.
XX		PR	26-SEP-2000;	2000US-0235484.
OS	Homo sapiens.	PR	27-SEP-2000;	2000US-0235834.
XX		PR	27-SEP-2000;	2000US-0235836.
PN	WO20015301-A2.	PR	29-SEP-2000;	2000US-0236327.
XX		PR	29-SEP-2000;	2000US-0236367.
PD		PR	29-SEP-2000;	2000US-0236368.
XX		PR	29-SEP-2000;	2000US-0236369.
XX		PR	29-SEP-2000;	2000US-0236370.
PF	17-JAN-2001; 2001WO-US01239.	PR	02-OCT-2000;	2000US-0236802.
XX		PR	02-OCT-2000;	2000US-0237037.
XX		PR	02-OCT-2000;	2000US-0237038.
XX		PR	02-OCT-2000;	2000US-0237039.
XX		PR	02-OCT-2000;	2000US-0237040.
XX		PR	13-OCT-2000;	2000US-0239335.
XX		PR	13-OCT-2000;	2000US-0239337.
XX		PR	20-OCT-2000;	2000US-0240960.
XX		PR	20-OCT-2000;	2000US-0241221.
XX		PR	20-OCT-2000;	2000US-0241785.
XX		PR	20-OCT-2000;	2000US-0241786.
XX		PR	20-OCT-2000;	2000US-0241787.
XX		PR	20-OCT-2000;	2000US-0241808.
XX		PR	20-OCT-2000;	2000US-0241809.
XX		PR	20-OCT-2000;	2000US-0241826.
XX		PR	01-NOV-2000;	2000US-0244617.
XX		PR	08-NOV-2000;	2000US-0246474.
XX		PR	08-NOV-2000;	2000US-0246475.
XX		PR	08-NOV-2000;	2000US-0246476.
XX		PR	08-NOV-2000;	2000US-0246477.
XX		PR	08-NOV-2000;	2000US-0246478.
XX		PR	08-NOV-2000;	2000US-0246523.
XX		PR	08-NOV-2000;	2000US-0246524.
XX		PR	08-NOV-2000;	2000US-0246525.
XX		PR	08-NOV-2000;	2000US-0246526.
XX		PR	08-NOV-2000;	2000US-0246527.
XX		PR	08-NOV-2000;	2000US-0246528.
XX		PR	08-NOV-2000;	2000US-0246532.
XX		PR	08-NOV-2000;	2000US-0246609.
XX		PR	08-NOV-2000;	2000US-0246610.
XX		PR	08-NOV-2000;	2000US-0246611.
XX		PR	08-NOV-2000;	2000US-0246613.
XX		PR	17-NOV-2000;	2000US-0249207.
XX		PR	17-NOV-2000;	2000US-0249208.
XX		PR	17-NOV-2000;	2000US-0249209.
XX		PR	17-NOV-2000;	2000US-0249210.
XX		PR	17-NOV-2000;	2000US-0249211.
XX		PR	17-NOV-2000;	2000US-0249212.
XX		PR	17-NOV-2000;	2000US-0249213.
XX		PR	17-NOV-2000;	2000US-0249214.
XX		PR	17-NOV-2000;	2000US-0249215.
XX		PR	17-NOV-2000;	2000US-0249216.
XX		PR	17-NOV-2000;	2000US-0249217.
XX		PR	17-NOV-2000;	2000US-0249218.
XX		PR	17-NOV-2000;	2000US-0249244.
XX		PR	17-NOV-2000;	2000US-0249245.
XX		PR	17-NOV-2000;	2000US-0249264.
XX		PR	17-NOV-2000;	2000US-0249265.
XX		PR	17-NOV-2000;	2000US-0249297.
XX		PR	17-NOV-2000;	2000US-0249299.
XX		PR	17-NOV-2000;	2000US-0249300.
XX		PR	01-DEC-2000;	2000US-0250160.
XX		PR	01-DEC-2000;	2000US-0250391.
XX		PR	05-DEC-2000;	2000US-0251030.
XX		PR	05-DEC-2000;	2000US-0251988.
XX		PR	05-DEC-2000;	2000US-0256719.
XX		PR	06-DEC-2000;	2000US-0251479.
XX		PR	08-DEC-2000;	2000US-0251856.
XX		PR	08-DEC-2000;	2000US-0251868.

PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX
 PS
 CC Claim 20; SEQ ID No 48529; 103pp; English.
 CC
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC disorders involving aberrant protein expression or biological activity.
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 796 AA;
 SQ
 Query Match 72.3%; Score 34; DB 22; Length 796;
 Best Local Similarity 75.0%; Pred. No. 4.1e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 TEGFFSGR 9
 DB 328 TEGFIAGR 335
 |||||:|

RESULT 13
 ABG28984
 ID ABG28984 standard; Protein; 896 AA.
 AC ABG28984;
 XX
 DT 18-FEB-2002 (first entry)
 DE
 XX Novel human diagnostic protein #28975.
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 KW
 XX Homo sapiens.
 OS
 XX WO200175067-A2.
 FN
 XX 11-OCT-2001.
 PD
 XX 30-MAR-2001; 2001WO-US08631.
 PF
 XX 31-MAR-2000; 2000US-0540217.
 PR
 XX 23-AUG-2000; 2000US-0649167.
 XX
 XX (HYSE-) HYSEQ INC.
 PA
 XX Drmanac RT, Liu C, Tang YT;
 PI
 XX WPI; 2001-639362/73.
 DR
 XX N-PSDB; AAS93171.
 DR
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -

XX
 PS
 CC Claim 20; SEQ ID No 59343; 103pp; English.
 CC
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC disorders involving aberrant protein expression or biological activity.
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 896 AA;

Query Match 72.3%; Score 34; DB 22; Length 896;
 Best Local Similarity 75.0%; Pred. No. 4.6e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 TEGFFSGR 9
 DB 153 TEGFIAGR 160
 |||||:|

RESULT 14
 AAU52608
 ID AAU52608 standard; Protein; 66 AA.
 XX
 AC AAU52608;
 XX
 DT 27-FEB-2002 (first entry)
 DE
 XX Propionibacterium acnes immunogenic protein #13504.
 XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 XX
 XX WO200181581-A2.
 FN
 XX 01-NOV-2001.
 PD
 XX 20-APR-2001; 2001WO-US12865.
 PF
 XX 21-APR-2000; 2000US-199047P.
 PR
 XX 02-JUN-2000; 2000US-206841P.
 PR
 XX 07-JUL-2000; 2000US-216747P.
 XX
 XX (CORI-) CORIXA CORP.
 PA
 XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 PI
 XX WPI; 2001-616774/71.
 DR
 XX N-PSDB; AAS9555.
 DR
 XX Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for

```

RESULT 2
Q8WU84      PRELIMINARY;      PRT; 1020 AA.
ID Q8WU84;
AC Q8WU84;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC021123; AAH21123.1; -.
DR InterPro; IPR001715; Calponin-like.
DR Pfam; PF00307; CH; 1.
DR SMART; SM00033; CH; 1.
DR PROSITE; PS50021; CH; 1.
KW Hypothetical protein.
FT NON_CODING_1
SQ SEQUENCE 1020 AA; 113890 MW; 1F1D124CDF6AC8A3 CRC64;

Query Match      83.3%; Score 35; DB 4; Length 1020;
Best Local Similarity 77.8%; Pred. No. 2e+02;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 STEVEAAGR 9
Db 744 SSEVDAAGR 752

RESULT 3
Q8ZNU1      PRELIMINARY;      PRT; 479 AA.
ID Q8ZNU1;
AC Q8ZNU1;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative cell wall-associated hydrolase.
GN STM1940.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan B., Sun H., Flores L., Miller W., Stoneking T., Nhan X.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
LT2."
RL Nature 413:852-856(2001).
RL EMBL; AB008786; AAL20855.1; -.
DR InterPro; IPR000064; NLPFC_P60.
DR Pfam; PF00877; NLPFC_P60; 1.
KW Hydrolase; Hypothetical protein; Complete proteome.
SQ SEQUENCE 479 AA; 54379 MW; DCCE378C23509013 CRC64;

Query Match      81.0%; Score 34; DB 16; Length 479;
Best Local Similarity 77.8%; Pred. No. 1.5e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 STEVEAAGR 9
Db 370 SAQVEAAGR 378

```

```

RESULT 4
Q8YM01      PRELIMINARY;      PRT; 110 AA.
ID Q8YM01;
AC Q8YM01;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein Alr5141.
GN Alr5141.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
cyanobacterium Anabaena sp. strain PCC 7120."
RL DNA Res. 8:205-213(2001)
DR EMBL; AP003598; BAB76840.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 110 AA; 11605 MW; B39BF5A6E6FD692A CRC64;

Query Match      78.6%; Score 33; DB 16; Length 110;
Best Local Similarity 87.5%; Pred. No. 54;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 STEVEAAG 8
Db 88 STEVETAG 95

RESULT 5
Q988D5      PRELIMINARY;      PRT; 223 AA.
ID Q988D5;
AC Q988D5;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Transcriptional regulator.
GN MLI6786.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
Mesorhizobium loti."
RL DNA Res. 7:331-338(2000).
RL EMBL; AP003010; BAB53015.1; -.
DR InterPro; IPR000524; HTH_GntR.
DR Pfam; PF00392; gntR; 1.
DR PRINTS; PR00035; HTHGNT.
DR SMART; SM00345; HTH_GNTR; 1.
DR PROSITE; PS00043; HTH_GNTR_FAMILY; 1.
KW Complete Proteome.
SQ SEQUENCE 223 AA; 24545 MW; 7984E6BBA63BEF56 CRC64;

Query Match      78.6%; Score 33; DB 16; Length 223;

```

DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE ATPase.
 GN MODC.
 OS Arthrobacter nicotinovorans.
 OG Plasmid pAO1.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Micrococcales; Micrococcaceae; Arthrobacter.
 OX NCBI_TaxID=29320;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9808982; PubMed=9428706;
 RA Menendez C., Otto A., Igloi G., Nick P., Brandsch R., Schubach B.,
 RA Boettcher B., Brandsch R.;
 RT "Molybdate-uptake genes and molybdopterin-biosynthesis genes on a
 RT bacterial plasmid. Characterization of MoeA as a filament-forming
 RT protein with adenosinetriphosphatase activity.";
 RL Eur. J. Biochem. 250:524-531(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Brandsch R.;
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBSJ databases.
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
 DR EMBL; Y10817; CAA1778.1; -.
 DR InterPro; IPR003593; AAA ATPase.
 DR InterPro; IPR003439; ABC transporter.
 DR InterPro; IPR005116; TOBE.
 DR Pfam; PF03459; TOBE; 1.
 DR ProDom; PD000006; ABC_transporter; 1.
 DR SMART; SM00382; AAA; 1.
 DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
 KW ATP-binding; Transport; Plasmid.
 SQ SEQUENCE 349 AA; 36687 MW; EBSA162121E39B39 CRC64;

 Query Match 78.6%; Score 33; DB 2; Length 349;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 2 TEVEAAG 8
 DB 121 TEVEAAG 127

 RESULT 10
 Q8GAG7
 ID Q8GAG7 PRELIMINARY; PRT; 349 AA.
 AC Q8GAG7
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Molybdenum transport ATPase modC.
 GN MODC.
 OS Arthrobacter nicotinovorans.
 OG Plasmid pAO1.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Micrococcales; Micrococcaceae; Arthrobacter.
 OX NCBI_TaxID=29320;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9511562; PubMed=7815950;
 RA Grether-Beck S., Igloi G.L., Fust S., Schiltz E., Decker K.,
 RA Brandsch R.;
 RT "Structural analysis and molybdenum-dependent expression of the pAO1-
 RT encoded nicotine dehydrogenase genes of Arthrobacter nicotinovorans.";
 RL Mol. Microbiol. 13:929-936(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96172783; PubMed=8588735;
 RA Menendez C., Igloi G., Henninger H., Brandsch R.;
 RA "A pAO1-encoded molybdopterin cofactor gene (moaA) of Arthrobacter
 RT nicotinovorans: characterization and site-directed mutagenesis of the

RT encoded protein.";
 RL Arch. Microbiol. 164:142-151(1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97230479; PubMed=9073580;
 RA Menendez C., Igloi G.L., Brandsch R.;
 RT "Is1473, a putative insertion sequence identified in the plasmid pAO1
 RT from Arthrobacter nicotinovorans: isolation, characterization and
 RT distribution among Arthrobacter species.";
 RL Plasmid 37:35-41(1997).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9808982; PubMed=9428706;
 RA Menendez C., Otto A., Igloi G., Nick P., Brandsch R., Schubach B.,
 RA Boettcher B., Brandsch R.;
 RT "Molybdate-uptake genes and molybdopterin-biosynthesis genes on a
 RT bacterial plasmid. Characterization of MoeA as a filament-forming
 RT protein with adenosinetriphosphatase activity.";
 RL Eur. J. Biochem. 250:524-531(1997).
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Schenk S., Hoelz A., Kraus B., Decker K.;
 RT "Gene structure and properties of enzymes of the plasmid-encoded
 RT nicotine catabolism of Arthrobacter nicotinovorans.";
 RL J. Mol. Biol. 284:1323-1339(1999).
 RN [6]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21405725; PubMed=11514508;
 RA Baitsch D., Sandu C., Brandsch R., Igloi G.L.;
 RT "A gene cluster on pAO1 of Arthrobacter nicotinovorans involved in the
 RT degradation of the plant alkaloid nicotine: cloning, purification and
 RT characterization of 2,6-dihydroxypyridine 3-hydroxylase.";
 RL J. Bacteriol. 183:5262-5267(2001).
 RN [7]
 RP SEQUENCE FROM N.A.
 RA Igloi G.L., Brandsch R.;
 RT "Sequence of the 165 Kb Catabolic Plasmid pAO1 from Arthrobacter
 RT nicotinovorans and Identification of a pAO1-dependent Nicotine Uptake
 RT System.";
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBSJ databases.
 KW EMBL; AJ507836; CAD47964.1; -.
 SQ SEQUENCE 349 AA; 36734 MW; 3817AE1CB261E280 CRC64;

 Query Match 78.6%; Score 33; DB 2; Length 349;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 2 TEVEAAG 8
 DB 121 TEVEAAG 127

 RESULT 11
 Q9A3Z5
 ID Q9A3Z5 PRELIMINARY; PRT; 738 AA.
 AC Q9A3Z5;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein CC3055.
 GN CC3055.
 OS Caulobacter crescentus.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
 OC Caulobacteraceae; Caulobacter.
 OX NCBI_TaxID=155892;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 19089 / CBL5;
 RX MEDLINE=21173698; PubMed=11259647;
 RA Niserman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,

```

DR InterPro; IPR006680; Amidohydro_1.
DR InterPro; IPR002195; Dihydroorotase.
DR InterPro; IPR005847; Urease.
DR Pfam; PF01979; Amidohydro_1; 1.
DR ProDom; PD000518; Urease; 1.
DR PROSITE; PS00483; DIHYDROOROTASE_2; 1.
KW Complete proteome.
SQ SEQUENCE 445 AA; 49295 MW; 24C2D0C0739CFC9F CRC64;

Query Match          76.2%; Score 32; DB 16; Length 445;
Best Local Similarity 66.7%; Pred. No. 3.5e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 STEVEAAGR 9
DB 42 SREIDAAGR 50

RESULT 15
Q9JJA1
ID Q9JJA1 PRELIMINARY; PRT; 459 AA.
AC Q9JJA1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Brain cDNA, clone MNCB-6015, similar to Mus musculus zinc finger
DE Protein 259 (Zfp259), mRNA.
DE Zfp259.
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL;
RA Osada N., Kusuda J., Tanuma R., Ito A., Hirata M., Sugano S.,
RT "isolation of full-length cDNA clones from mouse brain cDNA library
RT made by oligo-capping method.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR ENBL; AB041612; BAA95095.1; -.
DR MGD; MGI:1330262; Zfp259.
DR InterPro; IPR004457; Znf_ZPR1.
DR Pfam; PF03367; ZPR1; 2.
DR SMART; SMC0709; Zpr1; 2.
DR TIGRfam; TIGR00310; ZPR1 znf; 2.
SQ SEQUENCE 459 AA; 50741 MW; 771D38DD0806044F CRC64;

Query Match          76.2%; Score 32; DB 11; Length 459;
Best Local Similarity 55.6%; Pred. No. 3.7e+02;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 STEVEAAGR 9
DB 87 NTEIQSAGR 95

Search completed: December 11, 2003, 18:27:55
Job time : 25.3333 secs

```

DR InterPro; IPR002917; MMR_HSR1.
 DR InterPro; IPR005225; Small_GTP.
 DR InterPro; IPR004520; ThdF.
 DR Pfam; PF01926; MMR_HSR1; 1.
 DR PRINTS; PR00449; RASTRNSFRMG.
 DR TIGRFAMs; TIGR00650; MG442; 1.
 DR TIGRFAMs; TIGR00231; small_GTP; 1.
 DR TIGRFAMs; TIGR00450; thdF; 1.
 KW tRNA processing; GTP-binding; Complete proteome.
 FT NP_BIND 227 234 GTP (POTENTIAL).
 FT NP_BIND 274 278 GTP (POTENTIAL).
 FT NP_BIND 334 337 GTP (POTENTIAL).
 FT CONFLICT 264 264 L -> F (IN REF. 1).
 FT CONFLICT 302 302 I -> V (IN REF. 1).
 SQ SEQUENCE 456 AA; 51514 MW; 782753A7F29A21ED CRC64;
 Query Match 78.6%; Score 33; DB 1; Length 456;
 Best Local Similarity 87.5%; Pred. No. 19;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 TEVEAAGR 9
 DB 128 TEVEAAGR 135
 RESULT 2
 VCAP HSVB STANDARD; PRT; 1376 AA.
 AC P28920;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Major capsid protein (MCP) (Capsid protein VP5).
 GN 42.
 OS Equine herpesvirus type 1 (strain Ab4p) (EHV-1).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.
 OC NCBI_TaxID=31520;
 RN 1
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92295566; PubMed=1318606;
 RA Telford E.A.R., Watson M.S., McBride K., Davison A.J.;
 RT "The DNA sequence of equine herpesvirus-1.";
 RL Virology 189:304-316(1992).
 CC -!- FUNCTION: MAJOR PROTEIN OF THE ICOSAEDRAL CAPSID.
 CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES MAJOR CAPSID PROTEIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
 CC EMBL; M86664; AAB02478.1; -
 CC PIR; H36799; VCBED6.
 DR InterPro; IPR000912; Herpes_MCP.
 DR Pfam; PF03122; Herpes_MCP; 1.
 DR PRINTS; PR00235; HSVCAPSIDMCP.
 KW Coat protein.
 SQ SEQUENCE 1376 AA; 152182 MW; C3B866EFAE80AEDB CRC64;
 Query Match 78.6%; Score 33; DB 1; Length 1376;
 Best Local Similarity 87.5%; Pred. No. 56;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 TEVEAAGR 9
 DB 1280 TEVEAAGR 1287

RESULT 3
 ZPR1_HUMAN STANDARD; PRT; 459 AA.
 AC O75312;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Zinc-finger protein ZPR1 (Zinc finger protein 259).
 GN ZNF259 OR ZPR1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 RN 1
 RP SEQUENCE FROM N.A.
 RA Galcheva-Gargova Z., Konstantinov K.N., Gangwani L., Mikrut M., Purohit P., Theroux S.J., Enoch T., Davis R.J.;
 RT "Translational regulation by the ZPR1 signal transduction pathway.";
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
 RN 2
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain, Pancreas, and Skin;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.L., Wang J., Haieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A.C., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN 3
 RP SUBCELLULAR LOCATION.
 RX MEDLINE=98437195; PubMed=9763455;
 RA Galcheva-Gargova Z., Gangwani L., Konstantinov K.N., Mikrut M., Theroux S.J., Enoch T., Davis R.J.;
 RT "The cytoplasmic zinc finger protein ZPR1 accumulates in the nucleolus of proliferating cells.";
 RL Mol. Biol. Cell 9:2963-2971(1998).
 CC -!- FUNCTION: MAY BE A SIGNALING MOLECULE THAT COMMUNICATES MITOGENIC SIGNALS FROM THE CYTOPLASM TO THE NUCLEUS.
 CC -!- SUBUNIT: BINDS TO THE EGF AND PDGF RECEPTORS. BINDS TO THE ELONGATION FACTOR 1-ALPHA (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC; TRANSLOCATE TO THE NUCLEUS AFTER TREATMENT WITH MITOGENS.
 CC -!- SIMILARITY: BELONGS TO THE ZPR1 FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
 CC EMBL; AF019767; AAC33514.1; -
 CC EMBL; BC004256; AAH04256.1; -
 CC EMBL; BC012162; AAH12162.1; -
 CC EMBL; BC017349; AAH17349.1; -
 CC EMBL; BC017380; AAH17380.1; -
 CC Genbank; HGNC:13051; ZNF259.

core is deposited.

-1- MISCELLANEOUS: THERE ARE TWO TYPES OF FERRITIN SUBUNITS: L (LIGHT) CHAIN AND H (HEAVY) CHAIN. THE MAJOR CHAIN CAN BE LIGHT OR HEAVY, DEPENDING ON THE SPECIES AND TISSUE TYPE.

-1- SIMILARITY: BELONGS TO THE FERRITIN FAMILY.

-1- SIMILARITY: Contains 1 ferritin-like diiron domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; J04716; AAA37614.1; -
 DR EMBL; L39879; AAA62253.1; -
 DR EMBL; BC019840; AAH19840.1; -
 DR PIR; B33355; B33355.
 DR HSSP; P02791; 1DAT.
 DR SWISS-2DPAGE; P29391; MOUSE.
 DR MGD; MGI:95589; Ftl1.
 DR InterPro; IPR001519; Ferritin.
 DR Pfam; PF00210; ferritin; 1.
 DR ProDom; PD000971; Ferritin; 1.
 DR PROSITE; PS00540; FERRITIN_1; 1.
 DR PROSITE; PS00204; FERRITIN_2; 1.
 DR PROSITE; PS00905; FERRITIN-LIKE; 1.
 KW Iron storage; Iron; Metal-binding.
 FT INIT MET 0 0 BY SIMILARITY.
 FT DOMAIN 6 155 FERRITIN-LIKE DIIRON.
 FT METAL 53 53 IRON (POTENTIAL).
 FT METAL 56 56 IRON (POTENTIAL).
 FT METAL 57 57 IRON (POTENTIAL).
 FT METAL 60 60 IRON (POTENTIAL).
 FT METAL 63 63 IRON (POTENTIAL).
 FT METAL 63 63 IRON (POTENTIAL).
 FT CONFLICT 24 24 L -> V (IN REF. 3).
 FT CONFLICT 121 121 T -> A (IN REF. 3).
 SQ SEQUENCE 182 AA; 20671 MW; 2AC34371BAE27856 CRC64;

Query Match 73.8%; Score 31; DB 1; Length 182;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 STEVEAA 7
 Db |||||
 9 STEVEAA 15

RESULT 9
 FRL2_MOUSE STANDARD; PRT; 182 AA.
 AC P49945;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Ferritin light chain 2 (Ferritin L subunit 2) (Ferritin subunit LG).
 GN Ftl2 OR Ftl2-2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALE/c; TISSUE=Bone marrow;
 RX MEDLINE=92182532; PubMed=1543909;
 RA Renaudie F., Yachou A.K., Grandchamp B., Jones R., Beaumont C.;
 RT "A second ferritin L subunit is encoded by an intronless gene in the mouse."
 RL Mamm. Genome 2:143-149(1992).
 CC -1- FUNCTION: Ferritin is an intracellular molecule that stores iron in a soluble, nontoxic, readily available form. The functional molecule, which is composed of 24 chains, is roughly spherical and

contains a central cavity into which the polymeric ferric iron core is deposited.

-1- MISCELLANEOUS: THERE ARE TWO TYPES OF FERRITIN SUBUNITS: L (LIGHT) CHAIN AND H (HEAVY) CHAIN. THE MAJOR CHAIN CAN BE LIGHT OR HEAVY, DEPENDING ON THE SPECIES AND TISSUE TYPE.

-1- SIMILARITY: BELONGS TO THE FERRITIN FAMILY.

-1- SIMILARITY: Contains 1 ferritin-like diiron domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M73706; AAB00809.1; -
 DR HSSP; P02791; 1DAT.
 DR MGD; MGI:95590; Ftl2.
 DR InterPro; IPR001519; Ferritin.
 DR Pfam; PF00210; ferritin; 1.
 DR ProDom; PD000971; Ferritin; 1.
 DR PROSITE; PS00540; FERRITIN_1; 1.
 DR PROSITE; PS00204; FERRITIN_2; 1.
 DR PROSITE; PS00905; FERRITIN-LIKE; 1.
 KW Iron storage; Iron; Metal-binding.
 FT INIT MET 0 0 BY SIMILARITY.
 FT DOMAIN 6 155 FERRITIN-LIKE DIIRON.
 FT METAL 53 53 IRON (POTENTIAL).
 FT METAL 56 56 IRON (POTENTIAL).
 FT METAL 57 57 IRON (POTENTIAL).
 FT METAL 60 60 IRON (POTENTIAL).
 FT METAL 63 63 IRON (POTENTIAL).
 FT METAL 63 63 IRON (POTENTIAL).
 SQ SEQUENCE 182 AA; 20711 MW; 0F367FC59A47F00C CRC64;

Query Match 73.8%; Score 31; DB 1; Length 182;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 STEVEAA 7
 Db |||||
 9 STEVEAA 15

RESULT 10
 VGLI_HSVBS STANDARD; PRT; 380 AA.
 AC Q08102;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Glycoprotein I.
 GN GI.
 OS Bovine herpesvirus type 1.2 (strain ST).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.
 OX NCBI_TaxID=45407;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=94167875; PubMed=8122370;
 RA Leung-Tack P., Audonnet J.F., Riviere M.;
 RT "The complete DNA sequence and the genetic organization of the short unique region (US) of the bovine herpesvirus type 1 (ST strain).";
 RL Virology 199:409-421(1994).
 CC -1- SIMILARITY: TO OTHER HERPESVIRUSES GLYCOPROTEINS I, TO VZV GP1V, AND TO PRV GP63.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 60S acidic ribosomal protein P2B.
GN RPP2B.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-14.
RC STRAIN=cv. B73; TISSUE=Ear;
RX MEDLINE=97422884; PubMed=9276949;
RA Bailey-Serres J., Vancala S., Sick K., Lee C.H.;
RT "Acidic phosphoprotein complex of the 60S ribosomal subunit of maize
RL seedling roots. Components and changes in response to flooding.";
RL Plant Physiol. 114:11293-1305(1997).
CC -1- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE ELONGATION STEP OF
CC PROTEIN SYNTHESIS.
CC -1- SUBUNIT: P1 AND P2 EXIST AS DIMERS AT THE LARGE RIBOSOMAL
CC SUBUNIT (BY SIMILARITY).
CC -1- PTM: PHOSPHORYLATED.
CC -1- SIMILARITY: BELONGS TO THE L12P FAMILY OF RIBOSOMAL PROTEINS.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U62753; AAB71080.1; -.
CC PIR; T02040; T02040.
CC InterPro; IPR001813; 60S ribosomal.
CC Pfam; PF00428; 60S ribosomal; 1.
CC Ribosomal protein; Phosphorylation.
CC SEQUENCE 113 AA; 11700 MW; DCD7365951EA8ED7 CRC64;
Query Match 71.4%; Score 30; DB 1; Length 113;
Best Local Similarity 87.5%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 TEVEAGR 9
DB 52 TEVIAGR 59
RESULT 13
ID CBAA_COMTE STANDARD; PRT; 432 AA.
AC Q44256; Q08105;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 3-chlorobenzoate-3,4-dioxygenase oxygenase subunit (EC 1.14.-.-).
GN CBAA.
OS Comamonas testosteroni (Pseudomonas testosteroni).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Comamonadaceae; Comamonas.
OX NCBI_TaxID=285;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BR60 / Isolate Bloody Run creek; TRANSPOSON=tn5271;
RX MEDLINE=95219104; PubMed=7704279;
RA Nakatsu C.H., Straus N.A., Wyndham R.C.;
RT "The nucleotide sequence of the tn5271 3-chlorobenzoate
RT 3,4-dioxygenase genes (cbab) unites the class IA oxygenases in a
RT single lineage.";
RL Microbiology 141:485-495(1995).
CC -1- COFACTOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.
CC -1- SUBUNIT: THIS DIOXYGENASE SYSTEM CONSISTS OF TWO PROTEINS: AN
CC OXYGENASE AND AN OXYGENASE REDUCTASE.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING

```

```

CC DIOXYGENASE ALPHA SUBUNIT FAMILY.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U18133; AAC45716.1; -.
CC InterPro; IPR005806; Rieske dom.
CC InterPro; IPR001663; Ring_hydroxyl_A.
CC Pfam; PF00355; Rieske; 1.
CC PROSITE; PS00570; RING_HYDROXYL_ALPHA; FALSE NEG.
KW Aromatic hydrocarbons catabolism; Oxidoreductase; Metal-binding;
KW Iron-sulfur; Iron; 2Fe-2S; Dioxygenase; NAD.
FT METAL 69
FT METAL 71 71 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 88 88 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 91 91 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 180 180 IRON (BY SIMILARITY).
FT METAL 185 185 IRON (BY SIMILARITY).
SQ SEQUENCE 432 AA; 48927 MW; 216CF50FE14BEE2 CRC64;
Query Match 71.4%; Score 30; DB 1; Length 432;
Best Local Similarity 75.0%; Pred. No. 80;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 STEVEAG 8
DB 33 STELEAG 40
RESULT 14
SOXA_RHOSO
ID SOXA_RHOSO STANDARD; PRT; 453 AA.
AC P54995;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Dibenzothiophene desulfurization enzyme A.
GN SOXA OR DSZA.
OS Rhodococcus sp. (strain IGTS8).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Nocardiaceae; Rhodococcus.
OX NCBI_TaxID=1831;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=95050234; PubMed=7961424;
RA Denome S.A., Oldfield C., Nash L.J., Young K.D.;
RT "Characterization of the desulfurization genes from Rhodococcus sp.
RT strain IGTS8.";
RL J. Bacteriol. 176:6707-6716(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC MEDLINE=96031556; PubMed=7574582;
RA Piddington C.S., Kovacevich B.R., Rambossek J.;
RT "Sequence and molecular characterization of a DNA region encoding the
RT dibenzothiophene desulfurization operon of Rhodococcus sp. strain
RT IGTS8.";
RL Appl. Environ. Microbiol. 61:468-475(1995).
CC -1- FUNCTION: PART OF A PATHWAY TO REMOVE COVALENTLY BOUND SULFUR FROM
CC DIBENZOTHIOPHENE (DBT) WITHOUT BREAKING CARBON-CARBON BONDS. THIS
CC ENZYME METABOLIZES DBT-SULFONE (DBTOS OR DET 5,5-DIOXIDE) TO 2-
CC HYDROXYBIPHENYL (2-HBP).
CC -1- COFACTOR: FMN (POTENTIAL).
CC -1- PATHWAY: SECOND STEP IN PATHWAY FROM DBT TO 2-HBP.
CC -1- SUBUNIT: HETERODIMER OF TWO SUBUNITS, SOXA AND SOXB.
CC -1- SIMILARITY: BELONGS TO THE NTA/SNAA/SOXA (DSZA) FAMILY OF
CC MONOOXYGENASES.

```


Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TEVEAAG 8
DB 121 TEVEAAG 127

RESULT 3
T00126
hypothetical protein 4 - Leptospira interrogans
C:Species: Leptospira interrogans
C:Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 21-Jul-2000
C:Accession: T00126
R:Takahashi, Y.; Akase, K.; Hirano, H.; Fukunaga, M.
Gene 215, 37-45, 1998
A:Title: Physical and genetic maps of the Leptospira interrogans serovar icterohaemorrhagiae

A:Reference number: Z14115; MUID:98332717; PMID:9666070
A:Accession: T00126
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-435 <TAK>
A:Cross-references: EMBL:AB010203; NID:92780763; PIDN:BAA24373.1; PID:G2780770
A:Experimental source: strain Ictero No.1; substrain icterohaemorrhagiae
C:Superfamily: thiophen / furan oxidation protein; translation elongation factor Tu homolog

Query Match 78.6%; Score 33; DB 2; Length 435;
Best Local Similarity 87.5%; Pred. No. 29;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TEVEAAGR 9
DB 107 TEVEAAGR 114

RESULT 4
E87627
hypothetical protein CC3055 [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: E87627
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon, N.; J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: E87627
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-738 <STO>
A:Cross-references: GB:AE005673; NID:G13424701; PIDN:AAK25017.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC3055

Query Match 78.6%; Score 33; DB 2; Length 738;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EVEAAGR 9
DB 465 EVEAAGR 471

RESULT 5
VCBBD6
major capsid protein - equine herpesvirus 1 (strain Ab4p)
C:Species: equine herpesvirus 1
A:Note: host Equus caballus (domestic horse)
C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 16-Jul-1999
C:Accession: H36799
R:Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.

submitted to GenBank, March 1992
A:Description: The DNA sequence of equine herpesvirus-1.
A:Reference number: A36805
A:Accession: H36799
A:Molecule type: DNA
A:Residues: 1-1376 <TEL>
A:Cross-references: GB:M86664; NID:G330791; PIDN:AB02478.1; PID:G330835
R:Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.
Virology 189, 304-316, 1992
A:Title: The DNA sequence of equine herpesvirus-1.
A:Reference number: A41831; MUID:92295566; PMID:1318606
A:Contents: annotation; possible protein-coding frames
A:Note: neither amino acid nor nucleotide sequence is given
C:Genetics:
A:Gene: 42
C:Superfamily: varicella-zoster virus major capsid protein
C:Keywords: capsid protein

Query Match 78.6%; Score 33; DB 1; Length 1376;
Best Local Similarity 87.5%; Pred. No. 98;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TEVEAAGR 9
DB 1280 TEVEAAGR 1287

RESULT 6
B82954
probable dihydroorotase PA5541 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 10-May-2001
C:Accession: B82954
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Berman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, L.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen.
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: B82954
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-445 <STO>
A:Cross-references: GB:AE004966; GB:AE004091; NID:G9951872; PIDN:ARG08926.1; GSPDB:GN00148
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA5541
C:Superfamily: Bacillus dihydroorotase; Bacillus dihydroorotase homology

Query Match 76.2%; Score 32; DB 2; Length 445;
Best Local Similarity 66.7%; Pred. No. 49;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 STEVEAAGR 9
DB 42 SREIDAAGR 50

RESULT 7
S60792
M protein precursor - Streptococcus pyogenes (serotype M1) (fragment)
C:Species: Streptococcus pyogenes
A:Variety: serotype M1
C:Date: 20-Jul-1996 #sequence_revision 13-Mar-1997 #text_change 26-Aug-1999
C:Accession: S60792
R:Whitmore, A.M.; Kapur, V.; Sullivan, D.J.; Musser, J.M.; Kehoe, M.A.
Mol. Microbiol. 14, 619-631, 1994
A:Title: Non-congruent relationships between variation in emm gene sequences and the pC₁ gene.
A:Reference number: S60784; MUID:95198537; PMID:7891551
A:Accession: S60792
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-101 <WHA>

A;Cross-references: GB:J04716; NID:g193268; PIDN:AAA37614.1; PID:g309234
C;Superfamily: ferritin

Query Match 73.8%; Score 31; DB 2; Length 183;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 STEVEAA 7
|||||
Db 10 STEVEAA 16

RESULT 12

IS4774

ferritin light chain - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 20-Aug-1999

C;Accession: IS4774

R;Denis, M.G.

Int. J. Cancer 50, 930-936, 1992

A;Title: Isolation of cDNA clones corresponding to genes differentially expressed in two
A;Reference number: IS4774; MUID:92210224; PMID:1555892

A;Accession: IS4774

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-183 <RES>

A;Cross-references: GB:L01122; NID:g204122; PIDN:AAA41152.1; PID:g204123

C;Superfamily: ferritin

Query Match 73.8%; Score 31; DB 2; Length 183;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 STEVEAA 7
|||||
Db 10 STEVEAA 16

RESULT 13

AB2640

chemotaxis methyltransferase [imported] - Agrobacterium tumefaciens (strain C58, Dupont)

C;Species: Agrobacterium tumefaciens

C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002

C;Accession: AB2640

R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kuttyavin, T.; Levy, R.; Li, M.; McClell

; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,

ster, E.W.

A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A;Reference number: AB2577; MUID:21608550; PMID:11743193

A;Accession: AB2640

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-302 <KUR>

A;Cross-references: GB:AE008688; PIDN:AA141536.1; PID:g17738867; GSPDB:GN00186

A;Experimental source: strain C58 (Dupont)

C;Genetics:

A;Gene: CheR

A;Map position: circular chromosome

C;Superfamily: protein-glutamate O-methyltransferase; protein-glutamate O-methyltransfer

Query Match 73.8%; Score 31; DB 2; Length 302;
Best Local Similarity 75.0%; Pred. No. 54;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TEVEAAGR 9
|||||
Db 196 TEVDAGR 203

RESULT 14

C97422
cher homolog (AF044495) [imported] - Agrobacterium tumefaciens (strain C58, Cerson)
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002
C;Accession: C97422
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tu
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: C97422
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-302 <KUR>
A;Cross-references: GB:AE007869; PIDN:AAK86332.1; PID:g15155452; GSPDB:GN00169
C;Genetics:
A;Gene: AGR C.914
A;Map position: circular chromosome
C;Superfamily: protein-glutamate O-methyltransferase; protein-glutamate O-methyltransfe

Query Match 73.8%; Score 31; DB 2; Length 302;
Best Local Similarity 75.0%; Pred. No. 54;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TEVEAAGR 9
|||||
Db 196 TEVDAGR 203

RESULT 15

H87333

HLVD family secretion protein [imported] - Caulobacter crescentus

C;Species: Caulobacter crescentus

C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001

C;Accession: H87333

R;Niernman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A;Title: Complete Genome Sequence of Caulobacter crescentus.

A;Reference number: H87249; MUID:21173698; PMID:11259647

A;Accession: H87333

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-378 <STO>

A;Cross-references: GB:AE005673; NID:g13421902; PIDN:AAK22668.1; GSPDB:GN00148

C;Genetics:

A;Gene: CC0683

Query Match 73.8%; Score 31; DB 2; Length 378;
Best Local Similarity 66.7%; Pred. No. 69;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 STEVEAAGR 9
|||||
Db 217 ATQVEAAGR 225

Search completed: December 11, 2003, 18:29:17
Job time : 13.3333 secs

```
; FILE REFERENCE: P2040P1
; CURRENT APPLICATION NUMBER: US/10/042,141
; CURRENT FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: 09/726,643
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: PCT/US00/15187
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/137,725
; PRIOR FILING DATE: 1999-06-07
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 144
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-042-141-144

Query Match      83.3%; Score 35; DB 14; Length 42;
Best Local Similarity 77.8%; Pred. No. 4.4;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy      1 STEVEAAGR 9
Db      30 SSEVDAAGR 38

RESULT 3
US-09-726-643-143
; Sequence 143, Application US/09726643
; Patent No. US20020028449A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 26 Human secreted proteins
; FILE REFERENCE: P2040P1
; CURRENT APPLICATION NUMBER: US/09/726,643
; CURRENT FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: PCT/US00/15187
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/137,725
; PRIOR FILING DATE: 1999-06-07
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 143
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-726-643-143

Query Match      83.3%; Score 35; DB 9; Length 267;
Best Local Similarity 77.8%; Pred. No. 32;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy      1 STEVEAAGR 9
Db      30 SSEVDAAGR 38

RESULT 4
US-10-042-141-143
; Sequence 143, Application US/10042141
; Publication No. US20020183503A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 26 Human secreted proteins
; FILE REFERENCE: P2040P1
; CURRENT APPLICATION NUMBER: US/10/042,141
; CURRENT FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: 09/726,643
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: PCT/US00/15187
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/137,725
; PRIOR FILING DATE: 1999-06-07
```

```
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 143
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-042-141-143

Query Match      83.3%; Score 35; DB 14; Length 267;
Best Local Similarity 77.8%; Pred. No. 32;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy      1 STEVEAAGR 9
Db      30 SSEVDAAGR 38

RESULT 5
US-10-156-761-11994
; Sequence 11994, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11994
; LENGTH: 395
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-11994

Query Match      81.0%; Score 34; DB 15; Length 395;
Best Local Similarity 87.5%; Pred. No. 76;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      2 TEVEAAGR 9
Db      294 TEVKAAGR 301

RESULT 6
US-09-864-761-33858
; Sequence 33858, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
```

; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (59)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (124)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (247)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (270)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (282)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1260

Query Match 76.2%; Score 32; DB 10; Length 296;
Best Local Similarity 55.6%; Pred. No. 1.4e+02;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 STEVEAGR 9
Db 103 NTEIQSAGR 111

RESULT 10
US-10-289-757-80
; Sequence 80, Application US/10289757
; Publication No. US20030180751A1
; GENERAL INFORMATION:
; APPLICANT: Demmer, Jeroen
; APPLICANT: Forster, Richard L
; APPLICANT: Gibson, John Bryan
; APPLICANT: Shenk, Michael Andrew
; APPLICANT: No. US20030180751A1riss, Geoffrey
; APPLICANT: Glenn, Matthew
; APPLICANT: Saulsbury, Keith Martin
; APPLICANT: Hall, Claire
; TITLE OF INVENTION: Compositions isolated from forage
; FILE REFERENCE: 11000.1061U
; CURRENT APPLICATION NUMBER: US/10/289,757
; CURRENT FILING DATE: 2002-11-07
; PRIOR FILING DATE: 2002-11-07
; PRIOR FILING DATE: 2001-11-07
; NUMBER OF SEQ ID NOS: 218
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 80
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Lolium perenne
US-10-289-757-80

Query Match 76.2%; Score 32; DB 12; Length 390;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 STEVEAGR 9
Db 183 ATDVEAGR 191

RESULT 11
US-09-988-915-1
; Sequence 1, Application US/09988915
; Patent No. US20020102614A1
; GENERAL INFORMATION:
; APPLICANT: Davis, Roger J.
; APPLICANT: Gangwani, Laxman

; TITLE OF INVENTION: USE OF ZPR1 AS A MOLECULAR PROBE FOR SPINAL MUSCULAR ATROPHY
; FILE REFERENCE: 07917-132001
; CURRENT APPLICATION NUMBER: US/09/988,915
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: US 60/249,745
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Homo sapiens ZPR1
US-09-988-915-1

Query Match 76.2%; Score 32; DB 10; Length 459;
Best Local Similarity 55.6%; Pred. No. 2.3e+02;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 STEVEAGR 9
Db 87 NTEIQSAGR 95

RESULT 12
US-10-205-219-111
; Sequence 111, Application US/10205219
; Publication No. US20030138803A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; APPLICANT: Lee, Kevin
; APPLICANT: Dixon, Alistair
; APPLICANT: Brookbank, Robert
; APPLICANT: Pinmock, Robert
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
; FILE REFERENCE: WL-A-018200
; CURRENT APPLICATION NUMBER: US/10/205,219
; CURRENT FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: GB 0118354.0
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 111
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: Zinc finger protein ZPR1
US-10-205-219-111

Query Match 76.2%; Score 32; DB 12; Length 459;
Best Local Similarity 55.6%; Pred. No. 2.3e+02;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 STEVEAGR 9
Db 87 NTEIQSAGR 95

RESULT 13
US-10-289-757-172
; Sequence 172, Application US/10289757
; Publication No. US20030180751A1
; GENERAL INFORMATION:
; APPLICANT: Demmer, Jeroen
; APPLICANT: Forster, Richard L
; APPLICANT: Gibson, John Bryan
; APPLICANT: Shenk, Michael Andrew
; APPLICANT: No. US20030180751A1riss, Geoffrey
; APPLICANT: Glenn, Matthew
; APPLICANT: Saulsbury, Keith Martin
; APPLICANT: Hall, Claire
; TITLE OF INVENTION: Compositions isolated from forage
; APPLICANT: Gangwani, Laxman

```
DB          45 NTEIQSAGR 53

RESULT 2
US-08-870-518-1
; Sequence 1, Application US/08870518
; Patent No. 5925566
; GENERAL INFORMATION:
; APPLICANT: Davis, Roger J.
; APPLICANT: Galcheva-Gargova, Zoya
; TITLE OF INVENTION: NON-ACTIVATED RECEPTOR COMPLEX
; TITLE OF INVENTION: PROTEINS AND USES THEREOF
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/870,518
; FILING DATE: 06-JUN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/019,219
; FILING DATE: 06-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Fasse, Peter J.
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 04020/102001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 459 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-870-518-2

Query Match          76.2%; Score 32; DB 2; Length 459;
Best Local Similarity 55.6%; Pred. No. 46;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY          1 STEVEAAGR 9
           :|:|:|:|
DB          87 NTEIQSAGR 95

RESULT 4
US-09-252-991A-18889
; Sequence 18889, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18889
; LENGTH: 681
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-18889

Query Match          76.2%; Score 32; DB 4; Length 681;
Best Local Similarity 66.7%; Pred. No. 73;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY          1 STEVEAAGR 9
           :|:|:|:|
DB          278 SREIDAGR 286

RESULT 5
US-08-344-833-2
; Sequence 2, Application US/08344833
; Patent No. 5874280
; GENERAL INFORMATION:
```

```
; SEQ ID NO 18
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-761-962A-18

Query Match      71.4%; Score 30; DB 4; Length 359;
Best Local Similarity 75.0%; Pred. No. 95;
Matches 6; Conservative 1; Mismatches 0; Gaps 0;

QY 1 STEVEAAG 8
    |||:||||
Db 19 STELEGAG 26

RESULT 9
US-09-761-962A-21
; Sequence 21, Application US/09761962A
; Patent No. 6500927
; GENERAL INFORMATION:
; APPLICANT: Memorial Sloan-Kettering Cancer Center
; TITLE OF INVENTION: MULTIPLE SPICE VARIANTS OF THE MU-OPIOD RECEPTOR GENE
; FILE REFERENCE: 830002-2000.2
; CURRENT APPLICATION NUMBER: US/09/761,962A
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/743,872
; PRIOR APPLICATION NUMBER: PCT/US99/15974
; PRIOR FILING DATE: 1999-07-15
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 21
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-761-962A-21

Query Match      71.4%; Score 30; DB 4; Length 399;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 0; Gaps 0;

QY 1 STEVEAAG 8
    |||:||||
Db 19 STELEGAG 26

RESULT 10
US-09-252-991A-17259
; Sequence 17259, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17259
; LENGTH: 421
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17259

Query Match      71.4%; Score 30; DB 4; Length 421;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 0; Gaps 0;

QY 2 TEVEAAGR 9
    |||:||||
```

```
Db 124 TEVESLGR 131

RESULT 11
US-08-089-755A-2
; Sequence 2, Application US/08089755A
; Patent No. 5356801
; GENERAL INFORMATION:
; APPLICANT: Rambosek, John
; APPLICANT: Piddington, Chris
; APPLICANT: Kovacevich, Brian R
; APPLICANT: Young, Kevin D
; APPLICANT: Denome, Sylvia A
; TITLE OF INVENTION: Recombinant DNA Encoding A
; TITLE OF INVENTION: Desulfurization Biocatalyst
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith and Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/089,755A
; FILING DATE: 09-JUL-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/911,845
; FILING DATE: 10-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: EBC92-03A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 453 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-089-755A-2

Query Match      71.4%; Score 30; DB 1; Length 453;
Best Local Similarity 85.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 EVEAAGR 9
    |||:||||
Db 265 EVDAAGR 271

RESULT 12
US-08-421-754-2
; Sequence 2, Application US/08421754
; Patent No. 5578478
; GENERAL INFORMATION:
; APPLICANT: Rambosek, John
; APPLICANT: Piddington, Chris
; APPLICANT: Kovacevich, Brian R
; APPLICANT: Young, Kevin D
; APPLICANT: Denome, Sylvia A
; TITLE OF INVENTION: Recombinant DNA Encoding A
; TITLE OF INVENTION: Desulfurization Biocatalyst
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
```

NAME: Elmore, Carolyn S.
 REGISTRATION NUMBER: 37,567
 REFERENCE/DOCKET NUMBER: EBC96-06A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (781) 861-6240
 TELEFAX: (781) 861-9540
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 453 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-851-088-2

Query Match 71.4%; Score 30; DB 2; Length 453;
 Best Local Similarity 85.7%; Pred. No. 1.2e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 EYEAAGR 9
 |||||
 Db 265 EYDAAGR 271

RESULT 15

US-08-851-088-8
 ; Sequence 8, Application US/08951088
 ; Patent No. 5952208
 ; GENERAL INFORMATION:
 ; APPLICANT: Darzins, Aldis
 ; APPLICANT: Xi, Lei
 ; APPLICANT: Childs, John D.
 ; APPLICANT: Monticello, Daniel J.
 ; APPLICANT: Squires, Charles H.
 ; TITLE OF INVENTION: DSZ Gene Expression In Pseudomonas Hosts
 ; NUMBER OF SEQUENCES: 23
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
 ; STREET: Two Militia Drive
 ; CITY: Lexington
 ; STATE: Massachusetts
 ; COUNTRY: USA
 ; ZIP: 02173

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/851,088
 FILING DATE: 05-MAY-1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/835,185
 FILING DATE: 07-APR-1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Elmore, Carolyn S.
 REGISTRATION NUMBER: 37,567
 REFERENCE/DOCKET NUMBER: EBC96-06A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (781) 861-6240
 TELEFAX: (781) 861-9540
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 453 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-851-088-8

Query Match 71.4%; Score 30; DB 2; Length 453;
 Best Local Similarity 85.7%; Pred. No. 1.2e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 EYEAAGR 9
 |||||
 Db 265 QVEAAGR 271
 Search completed: December 11, 2003, 18:30:32
 Job time : 10.2222 secs

```

XX 12-MAY-1998; 98US-0085197.
XX 05-MAY-1999; 99US-0085197.
XX (BOCK/) BOCK S C.
XX (PICA/) PICARD V.
XX (ZEND/) ZENDEHROUH P.
XX Bock SC, Picard V, Zendehtrouh P;
XX WPI; 2000-116274/10.
XX
XX New modified human antithrombin III compounds, used for treating e.g.
XX sepsis, trauma, acute respiratory distress syndrome, restenosis,
XX thrombosis, thromboembolism or stroke -
XX
XX Claim 13; Page 57; 75pp; English.
XX
XX The present sequence is from an antithrombin III (ATIII) variant, 7BEVA
XX derived from human ATIII.N135A cDNA insert of the pBluebac baculovirus
XX expression construct and comprises residues 385-393. The variant has
XX improved resistance to elastase and IGG-activated neutrophils while
XX retains anti-thrombin and anti-factor Xa activities. It may be
XX expressed as glycoforms with enhanced heparin affinity which target the
XX blood vessel wall more efficiently than ATIIIs with normal heparin
XX affinity. The modified ATIIIs can be used to treat thrombin activation-
XX related pathological symptoms due to sepsis, trauma, acute
XX respiratory distress syndrome, restenosis, thrombosis, thromboembolism
XX and stroke. It can also be used to reduce the risk of reocclusion
XX and restenosis in percutaneous transluminal coronary angioplasty,
XX and thrombosis associated with surgery, ischaemia/reperfusion injury, and
XX coagulation abnormalities in cancer or surgical patients.
XX
XX Sequence 9 AA;
XX
XX Query Match 100.0%; Score 42; DB 21; Length 9;
XX Best Local Similarity 100.0%; Pred. No. 9.3e+05;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 STEVEAAGR 9
XX |||||
XX Db 1 STEVEAAGR 9
XX
XX RESULT 2
XX AAY44474
XX AC AAY44474;
XX
XX 27-MAR-2000 (first entry)
XX
XX Human antithrombin III variant 13.B (residues 385-393).
XX
XX Human; antithrombin III; ATIII variant 13.B; elastase-resistant;
XX IGG activated neutrophil resistant; anti-thrombin activity; heparin;
XX anti-factor Xa activity; blood clotting disorder; sepsis; trauma; stroke;
XX thrombin activation-related pathological symptom; restenosis; thrombosis;
XX acute respiratory distress syndrome; thromboembolism; reocclusion.
XX
XX Homo sapiens.
XX OS Synthetic.
XX
XX Key Location/Qualifiers
XX Key Misc-difference 3
XX FT Misc-difference 5 /note= "ATIII.N135A Ala at 387 is substituted by Glu"
XX FT Misc-difference 5 /note= "ATIII.N135A Val at 389 is substituted by Glu"
XX FT Misc-difference 6 /note= "ATIII.N135A Ile at 390 is substituted by Gly"
XX FT Misc-difference 6 /note= "ATIII.N135A Ile at 390 is substituted by Gly"
XX
XX WO9956098-A2.
XX

```

```

PD 18-NOV-1999.
XX
XX 12-MAY-1999; 99WO-US10549.
XX
XX 12-MAY-1998; 98US-0085197.
XX 05-MAY-1999; 99US-0085197.
XX
XX (BOCK/) BOCK S C.
XX (PICA/) PICARD V.
XX (ZEND/) ZENDEHROUH P.
XX
XX Bock SC, Picard V, Zendehtrouh P;
XX WPI; 2000-116274/10.
XX
XX New modified human antithrombin III compounds, used for treating e.g.
XX sepsis, trauma, acute respiratory distress syndrome, restenosis,
XX thrombosis, thromboembolism or stroke -
XX
XX Claim 13; Page 57; 75pp; English.
XX
XX The present sequence is from an antithrombin III (ATIII) variant, 13.B
XX derived from human ATIII.N135A cDNA insert of the pBluebac baculovirus
XX expression construct and comprises residues 385-393. The variant has
XX improved resistance to elastase and IGG-activated neutrophils while
XX retains anti-thrombin and anti-factor Xa activities. It may be
XX expressed as glycoforms with enhanced heparin affinity which target the
XX blood vessel wall more efficiently than ATIIIs with normal heparin
XX affinity. The modified ATIIIs can be used to treat thrombin activation-
XX related pathological symptoms due to sepsis, trauma, acute
XX respiratory distress syndrome, restenosis, thrombosis, thromboembolism
XX and stroke. It can also be used to reduce the risk of reocclusion
XX and restenosis in percutaneous transluminal coronary angioplasty,
XX and thrombosis associated with surgery, ischaemia/reperfusion injury, and
XX coagulation abnormalities in cancer or surgical patients.
XX
XX Sequence 9 AA;
XX
XX Query Match 90.5%; Score 38; DB 21; Length 9;
XX Best Local Similarity 88.9%; Pred. No. 9.3e+05;
XX Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 STEVEAAGR 9
XX |||||
XX Db 1 STEVEAAGR 9
XX
XX RESULT 3
XX AAY44467
XX ID AAY44467 standard; peptide; 9 AA.
XX
XX AC AAY44467;
XX
XX 27-MAR-2000 (first entry)
XX
XX Human antithrombin III variant 5EA (residues 385-393).
XX
XX Human; antithrombin III; ATIII variant 5EA; elastase-resistant;
XX IGG activated neutrophil resistant; anti-thrombin activity; heparin;
XX anti-factor Xa activity; blood clotting disorder; sepsis; trauma; stroke;
XX thrombin activation-related pathological symptom; restenosis; thrombosis;
XX acute respiratory distress syndrome; thromboembolism; reocclusion.
XX
XX Homo sapiens.
XX OS Synthetic.
XX
XX Key Location/Qualifiers
XX Key Misc-difference 5 /note= "ATIII.N135A Val at 389 is substituted by Glu"
XX FT Misc-difference 6 /note= "ATIII.N135A Ile at 390 is substituted by Ala"
XX FT Misc-difference 6 /note= "ATIII.N135A Ile at 390 is substituted by Ala"
XX
XX WO9956098-A2.
XX

```



```

PF 02-JUN-2000; 2000WO-US15187.
XX
XX
PR 07-JUN-1999; 99US-0137725.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ruben SM, Birse CE, Duan RD, Soppet DR, Rosen CA, Shi Y;
PI Lafleur DW, Olsen HS, Ebner R, Florence KA, Ni J, Young PS;
XX WPI; 2001-061741/07.
XX
XX Nucleic acids encoding 26 human secreted polypeptides, useful e.g. for
PT preventing, diagnosing and/or treating cancers and for promoting wound
PT healing -
XX
XX Disclosure; Page 61; 530pp; English.
XX
XX The present invention relates to 26 secreted human proteins. The
CC proteins may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate polypeptide expression.
CC For example, they may be used in gene therapy or in vaccines.
CC Typical of diseases which are potentially treatable are cancers
CC (including leukemia), autoimmune diseases, allergies, inflammation,
CC graft rejection, hyperproliferation, cardiovascular diseases
CC (particularly critical limb ischemia and coronary disease) and any
CC involving abnormal angiogenesis, neurodegeneration and/or
XX infectious diseases.
XX
XX Sequence 42 AA;
SQ
Query Match 83.3%; Score 35; DB 22; Length 42;
Best Local Similarity 77.8%; Pred. No. 2.1;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
OY 1 STEVERAAGR 9
Db |::|::|
30 SSEVDAAGR 38
RESULT 6
AAB65084
ID AAB65084 standard; peptide; 267 AA.
XX
XX AAB65084;
XX
XX 23-MAR-2001 (first entry)
XX
XX Gene #19 associated peptide #2.
XX
XX Secreted protein; gene therapy; vaccine; cancer; leukemia;
KW autoimmune disease; allergy; inflammation; graft rejection;
KW hyperproliferation; cardiovascular; infection.
XX
XX Homo sapiens.
XX
XX WO2000075375-A1.
XX
XX 14-DEC-2000.
XX
XX 02-JUN-2000; 2000WO-US15187.
XX
XX 07-JUN-1999; 99US-0137725.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ruben SM, Birse CE, Duan RD, Soppet DR, Rosen CA, Shi Y;
PI Lafleur DW, Olsen HS, Ebner R, Florence KA, Ni J, Young PS;
XX WPI; 2001-061741/07.
XX
XX Nucleic acids encoding 26 human secreted polypeptides, useful e.g. for
PT preventing, diagnosing and/or treating cancers and for promoting wound
PT healing -
XX
XX Disclosure; Page 61; 530pp; English.
XX
XX The present invention relates to 26 secreted human proteins. The
CC proteins may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate polypeptide expression.
CC For example, they may be used in gene therapy or in vaccines.
CC Typical of diseases which are potentially treatable are cancers
CC (including leukemia), autoimmune diseases, allergies, inflammation,
CC graft rejection, hyperproliferation, cardiovascular diseases
CC (particularly critical limb ischemia and coronary disease) and any
CC involving abnormal angiogenesis, neurodegeneration and/or
XX infectious diseases.
XX
XX Sequence 42 AA;
SQ
Query Match 83.3%; Score 35; DB 22; Length 42;
Best Local Similarity 77.8%; Pred. No. 2.1;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
OY 1 STEVERAAGR 9
Db |::|::|
30 SSEVDAAGR 38
RESULT 6
AAB65084
ID AAB65084 standard; peptide; 267 AA.
XX
XX AAB65084;
XX
XX 23-MAR-2001 (first entry)
XX
XX Gene #19 associated peptide #2.
XX
XX Secreted protein; gene therapy; vaccine; cancer; leukemia;
KW autoimmune disease; allergy; inflammation; graft rejection;
KW hyperproliferation; cardiovascular; infection.
XX
XX Homo sapiens.
XX
XX WO2000075375-A1.
XX
XX 14-DEC-2000.
XX
XX 02-JUN-2000; 2000WO-US15187.
XX
XX 07-JUN-1999; 99US-0137725.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ruben SM, Birse CE, Duan RD, Soppet DR, Rosen CA, Shi Y;
PI Lafleur DW, Olsen HS, Ebner R, Florence KA, Ni J, Young PS;
XX WPI; 2001-061741/07.
XX
XX Nucleic acids encoding 26 human secreted polypeptides, useful e.g. for
PT preventing, diagnosing and/or treating cancers and for promoting wound
PT healing -
XX
XX Disclosure; Page 61; 530pp; English.
XX
XX The present invention relates to 26 secreted human proteins. The
CC proteins may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate polypeptide expression.
CC For example, they may be used in gene therapy or in vaccines.
CC Typical of diseases which are potentially treatable are cancers
CC (including leukemia), autoimmune diseases, allergies, inflammation,
CC graft rejection, hyperproliferation, cardiovascular diseases
CC (particularly critical limb ischemia and coronary disease) and any
CC involving abnormal angiogenesis, neurodegeneration and/or
XX infectious diseases.
XX
XX Sequence 267 AA;
SQ
Query Match 83.3%; Score 35; DB 22; Length 267;
Best Local Similarity 77.8%; Pred. No. 19;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
OY 1 STEVERAAGR 9
Db |::|::|
30 SSEVDAAGR 38
RESULT 7
ABP78180
ID ABP78180 standard; Protein; 66 AA.
XX
XX ABP78180;
XX
XX 07-MAR-2003 (first entry)
XX
XX N. gonorrhoeae amino acid sequence SEQ ID 2890.
XX
XX Antibacterial; infection; vaccine; gene therapy.
XX
XX Neisseria gonorrhoeae.
XX
XX WO200279243-A2.
XX
XX 10-OCT-2002.
XX
XX 12-FEB-2002; 2002WO-IB02069.
XX
XX 12-FEB-2001; 2001GB-0003424.
XX
XX (CHIR-) CHIRON SPA.
XX
XX Fontana MR, Pizsa M, Masignani V, Monaci E;
XX WPI; 2003-058415/05.
XX
XX N-PSDE; ABZ39150.
XX
XX New protein from Neisseria gonorrhoeae, useful for the manufacture of a
PT medicament for treating or preventing N. gonorrhoeae infection -
XX
XX Disclosure; Page 400; 815pp; English.
XX
XX The present invention relates to proteins from Neisseria gonorrhoeae.
CC Also disclosed are the nucleic acid molecules encoding the proteins and
CC antibodies that specifically bind to the proteins. The composition
CC comprising the protein, nucleic acid or antibody is useful for the
CC manufacture of a medicament for treating or preventing N. gonorrhoeae
CC infection, this may be in the form of a vaccine or gene therapy.
CC Sequences given in records ABP76736-ABP81046 represent nucleic acid
CC molecules of the invention.
XX
XX Sequence 66 AA;
SQ
Query Match 81.0%; Score 34; DB 24; Length 66;
Best Local Similarity 87.5%; Pred. No. 6;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

XX Human, antithrombin III; ATIII variant Bb.A; elastase-resistant;
 KW IGG activated neutrophil resistant; anti-thrombin activity; heparin;
 KW anti-factor Xa activity; blood clotting disorder; sepsis; trauma; stroke;
 KW thrombin activation-related pathological symptom; restenosis; thrombosis;
 KW acute respiratory distress syndrome; thromboembolism; reocclusion.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FH Misc-difference 3 /note= "ATIII.N135A Ala at 387 is substituted by Glu"
 FT Misc-difference 4 /note= "ATIII.N135A Val at 388 is substituted by Gly"
 FT Misc-difference 5 /note= "ATIII.N135A Val at 389 is substituted by Glu"
 FT Misc-difference 6 /note= "ATIII.N135A Val at 390 is substituted by Ala"
 FT Misc-difference 7 /note= "ATIII.N135A Ala at 391 is substituted by Ser"
 FT
 XX WO9958098-A2.
 XX 18-NOV-1999.
 XX 12-MAY-1999; 99WO-US10549.
 XX 12-MAY-1998; 98US-0085197.
 XX 05-MAY-1999; 99US-0085197.
 XX (BOCK/) BOCK S C.
 PA (PICA/) PICARD V.
 PA (ZEND/) ZENDEHROUH P.
 XX
 PI Bock SC, Picard V, Zendehehrouh P;
 XX
 XX WPI; 2000-116274/10.
 XX
 XX New modified human antithrombin III compounds, used for treating e.g.
 PT sepsis, trauma, acute respiratory distress syndrome, restenosis,
 PT thrombosis, thromboembolism or stroke -
 XX
 PS Claim 13; Page 57; 75pp; English.
 XX
 XX The present sequence is from an antithrombin III (ATIII) variant, Bb.A
 CC derived from human ATIII.N135A cDNA insert of the pBlueBac baculovirus
 CC expression construct and comprises residues 385-393. The variant has
 CC improved resistance to elastase and IGG-activated neutrophils while
 CC retains anti-thrombin and anti-factor Xa activities. It may be
 CC expressed as glycoforms with enhanced heparin affinity which target the
 CC blood vessel wall more efficiently than ATIIIs with normal heparin
 CC affinity. The modified ATIIIs can be used to treat thrombin activation-
 CC related pathological symptoms due to sepsis, trauma, acute
 CC respiratory distress syndrome, restenosis, thrombosis, thromboembolism
 CC and stroke. It can also be used to reduce the risk of reocclusion
 CC and restenosis in percutaneous transluminal coronary angioplasty,
 CC thrombosis associated with surgery, ischaemia/reperfusion injury, and
 CC coagulation abnormalities in cancer or surgical patients.
 XX
 XX Sequence 9 AA;
 SQ
 Query Match 76.2%; Score 32; DB 21; Length 9;
 Best Local Similarity 77.8%; Pred. No. 9.3e+05;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 STEVEAGR 9
 ||| |||
 Db 1 STEGEASGR 9
 ||| |||
 RESULT 11
 AAY44473

ID AAY44473 standard; peptide; 9 AA.
 XX
 AC AAY44473;
 XX
 DT 27-MAR-2000 (first entry)
 XX
 DE Human antithrombin III variant 13.A (residues 385-393).
 XX
 XX Human; antithrombin III; ATIII variant 13.A; elastase-resistant;
 KW IGG activated neutrophil resistant; anti-thrombin activity; heparin;
 KW anti-factor Xa activity; blood clotting disorder; sepsis; trauma; stroke;
 KW thrombin activation-related pathological symptom; restenosis; thrombosis;
 KW acute respiratory distress syndrome; thromboembolism; reocclusion.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FH Misc-difference 5 /note= "ATIII.N135A Val at 389 is substituted by Glu"
 FT Misc-difference 6 /note= "ATIII.N135A Ile at 390 is substituted by Gly"
 FT
 XX WO9958098-A2.
 XX 18-NOV-1999.
 XX 12-MAY-1999; 99WO-US10549.
 XX 12-MAY-1998; 98US-0085197.
 XX 05-MAY-1999; 99US-0085197.
 XX (BOCK/) BOCK S C.
 PA (PICA/) PICARD V.
 PA (ZEND/) ZENDEHROUH P.
 XX
 PI Bock SC, Picard V, Zendehehrouh P;
 XX
 XX WPI; 2000-116274/10.
 XX
 XX New modified human antithrombin III compounds, used for treating e.g.
 PT sepsis, trauma, acute respiratory distress syndrome, restenosis,
 PT thrombosis, thromboembolism or stroke -
 XX
 PS Claim 13; Page 57; 75pp; English.
 XX
 XX The present sequence is from an antithrombin III (ATIII) variant, 13.A
 CC derived from human ATIII.N135A cDNA insert of the pBlueBac baculovirus
 CC expression construct and comprises residues 385-393. The variant has
 CC improved resistance to elastase and IGG-activated neutrophils while
 CC retains anti-thrombin and anti-factor Xa activities. It may be
 CC expressed as glycoforms with enhanced heparin affinity which target the
 CC blood vessel wall more efficiently than ATIIIs with normal heparin
 CC affinity. The modified ATIIIs can be used to treat thrombin activation-
 CC related pathological symptoms due to sepsis, trauma, acute
 CC respiratory distress syndrome, restenosis, thrombosis, thromboembolism
 CC and stroke. It can also be used to reduce the risk of reocclusion
 CC and restenosis in percutaneous transluminal coronary angioplasty,
 CC thrombosis associated with surgery, ischaemia/reperfusion injury, and
 CC coagulation abnormalities in cancer or surgical patients.
 XX
 XX Sequence 9 AA;
 SQ
 Query Match 76.2%; Score 32; DB 21; Length 9;
 Best Local Similarity 77.8%; Pred. No. 9.3e+05;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 STEVEAGR 9
 ||| |||
 Db 1 STAVEGAGR 9
 ||| |||
 RESULT 12

```

XX Human; foetal liver; gene expression; single exon nucleic acid probe.
XX Homo sapiens.
XX WO200157277-A2.
XX 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US00669.
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483447/52.
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human fetal liver -
XX Claim 27; SEQ ID NO 25727; 639pp + sequence listing; English.
XX The invention relates to a single exon nucleic acid probe for
XX measuring human gene expression in a sample derived from human foetal
XX liver. The single exon nucleic acid probes may be used for predicting,
XX measuring and displaying gene expression in samples derived from human
XX fetal liver. The present sequence is a peptide encoded by a single exon
XX nucleic acid probe of the invention.
XX Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 54 AA;
XX Query Match 76.2%; Score 32; DB 22; Length 54;
XX Best Local Similarity 55.6%; Pred. No. 13;
XX Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
XX QY 1 STEVEAAGR 9
XX Db :||:||||
XX 30 NTEIQSAGR 38
XX RESULT 15
XX ID ABB18560 standard; Protein; 54 AA.
XX AC ABB18560;
XX XX
XX 23-JAN-2002 (first entry)
XX DE Protein #559 encoded by probe for measuring heart cell gene expression.
XX Human; gene expression; heart; microarray; vascular system;
XX cardiovascular disease; hypertension; cardiac arrhythmia;
XX congenital heart disease.
XX Homo sapiens.
XX WO200157274-A2.
XX 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US00666.
XX
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488899/53.
XX Single exon nucleic acid probes for analyzing gene expression in human
XX hearts -
XX Claim 15; SEQ ID NO 20330; 530pp; English.
XX The present invention relates to single exon nucleic acid probes for
XX measuring human gene expression in a sample derived from human heart (see
XX AB21535-AB41305). The present sequence is a protein encoded by one such
XX probe. The probes may be used for predicting, measuring and displaying
XX gene expression in samples derived from the human heart via microarrays.
XX By measuring gene expression, the probes are useful for predicting,
XX diagnosing, grading, staging, monitoring and prognosing diseases of the
XX human heart and vascular system e.g. cardiovascular disease,
XX hypertension, cardiac arrhythmias and congenital heart disease.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 54 AA;
XX Query Match 76.2%; Score 32; DB 22; Length 54;
XX Best Local Similarity 55.6%; Pred. No. 13;
XX Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
XX QY 1 STEVEAAGR 9
XX Db :||:||||
XX 30 NTEIQSAGR 38
XX Search completed: December 11, 2003, 18:24:37
XX Job time : 27 secs

```

```

RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities";
RL Nature 417:459-463 (2002).
DR EMBL; AE011673; AAM35356.1; -.
DR InterPro; IPR001570; Peptidase_M4.
DR InterPro; IPR005075; Pep_M4_propep.
DR Pfam; PF01447; Peptidase_M4; 1.
DR Pfam; PF02666; Peptidase_M4_C1; 1.
DR Pfam; PF03413; Pep_M4_propep; 1.
DR PRINTS; PR00730; THERMOLYSIN.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Complete proteome.
SQ SEQUENCE 540 AA; 57807 MW; F9A5737215D34162 CRC64;

Query Match 85.4%; Score 35; DB 16; Length 540;
Best Local Similarity 88.9%; Pred.No. 56;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps

QY 1 STALEAIGR 9
DB 476 STALTAIGR 484

RESULT 4
O69124 PRELIMINARY; PRT; 307 AA.
ID O69124
AC O69124;
CT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative glycosyl transferase WBIC.
GE WBIC.
OS Burkholderia pseudomallei (Pseudomonas pseudomallei), and
OS Burkholderia mallei (Pseudomonas mallei).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=28450; 13373;
[1]_RN
RN SEQUENCE FROM N.A.
RP SPECIES=B.pseudomallei; STRAIN=1026B;
RC Deshazer D., Brett P.J., Woods D.E.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
[2]_RN
RN SEQUENCE FROM N.A.
RP SPECIES=B.mallei;
RC Burnick M.N., Brett P.J., Woods D.E.;
RT "Physical and Molecular Characterization of Lipopolysaccharide O-
RT antigens from Burkholderia mallei."
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF064070; AA05462.1; -.
DR EMBL; AY028370; AA027399.1; -.
DR InterPro; IPR001173; Glyco_trans_2.
KW Transferase.
SQ SEQUENCE 307 AA; 33552 MW; E7964082B8C5BDE3 CRC64;

Query Match 82.9%; Score 34; DB 2; Length 307;
Best Local Similarity 87.5%; Pred.No. 50;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps

QY 2 TALEAIGR 9
DB 30 TALESIGR 37

RESULT 5
Q90WY7 PRELIMINARY; PRT; 71 AA.
ID Q90WY7
AC Q90WY7;
CT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-CCT-2002 (TrEMBLrel. 22, Last annotation update)

```

Query Match

```

RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=83241725; PubMed=6864790;
RA  Dunn J.J., Studier F.W.;
RT  "Complete nucleotide sequence of bacteriophage T7 DNA and the
RL  locations of T7 genetic elements.";
RJ  J. Mol. Biol. 166:477-535(1993).
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation
CC  at the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; V01146; CAA24425.1; -.
DR  PIR; A04354; JQBPT7.
DR  Structural protein.
KW  SEQUENCE 536 AA; 59120 MW; CDE87B92DC4A6C65 CRC64;
SQ  SEQUENCE 536 AA; 59120 MW; CDE87B92DC4A6C65 CRC64;

Query Match 90.2%; Score 37; DB 1; Length 536;
Best Local Similarity 88.9%; Pred.No. 2.4;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY  1 STALEAIGR 9
DB  427 STGLEAIGR 435

RESULT 3
ID  YE96 MYCTU STANDARD; PRT; 334 AA.
AC  P1177;
DT  15-DEC-1998 (Rel. 37, Created)
DT  15-DEC-1998 (Rel. 37, Last sequence update)
DT  15-SEP-2003 (Rel. 42, Last annotation update)
DE  Hypothetical protein Rv1496.
OS  Rv1496 OR MT1543 OR MTCY277.18.
GN  Mycobacterium tuberculosis.
OC  Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC  Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX  NCBI_TaxID=1773;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  STRAIN=H37RV;
RX  MEDLINE=98295987; PubMed=9634230;
RA  Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA  Gordon S.V., Eiglmeyer K., Gas S., Barry C.E. III, Tekai F.,
RA  Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA  Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA  Hornsby T., Jagals K., Krogh A., McLean J., Moule S., Murphy L.,
RA  Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA  Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA  Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT  "Deciphering the biology of Mycobacterium tuberculosis from the
RT  complete genome sequence.";
RL  Nature 393:537-544(1998).
RN  [2]
RP  SEQUENCE FROM N.A.
RX  STRAIN=CDC 1551 / Oshkosh;
RA  Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA  Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA  Kolonay J.F., Nelson W.C., Unayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA  Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA  Bishai W.;
RT  "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT  laboratory strains.";
RL  Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC  -!- SIMILARITY: BELONGS TO THE ARKG FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; Z79701; CAB02046.1; -.
DR  EMBL; AE007022; AAK45810.1; -.
DR  PIR; C70712; C70712.
DR  TIGR; MT1543; -.
DR  TubercuList; Rv1496; -.
DR  InterPro; IPR005129; ArgK.
DR  Pfam; PF03308; ArgK; 1.
DR  TIGRFAMs; TIGR00750; lso; 1.
KW  Hypothetical protein; ATP-Binding; Complete proteome.
FT  NP_BIND 65 72 ATP (POTENTIAL).
SQ  SEQUENCE 334 AA; 36256 MW; FFE57F8C7E6C38B5 CRC64;

Query Match 78.0%; Score 32; DB 1; Length 334;
Best Local Similarity 75.0%; Pred.No. 18;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY  1 STALEAIG 8
DB  72 STALEALG 79

RESULT 4
ID  TERM_ADEG1 STANDARD; PRT; 575 AA.
AC  Q64752;
DT  01-NOV-1997 (Rel. 35, Created)
DT  01-NOV-1997 (Rel. 35, Last sequence update)
DT  01-NOV-1997 (Rel. 35, Last annotation update)
DE  DNA terminal protein (Bellett protein) (pTP protein).
GN  pTP.
OS  Avian adenovirus gall (strain Phelps) (Fowl adenovirus 1) (CELO).
OC  Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Aviadenovirus.
OX  NCBI_TaxID=10553;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=96186720; PubMed=8627769;
RA  Chlocca S., Kurzbauer R., Schaffner G., Baker A., Mautner V.,
RA  Cotten M.;
RT  "The complete DNA sequence and genomic organization of the avian
RT  adenovirus CELO.";
RL  J. Virol. 70:2939-2949(1996).
CC  -!- FUNCTION: THIS PROTEIN IS COVALENTLY ATTACHED TO THE TERMINI OF
CC  REPLICATING DNA IN VIVO & NASCENT DNA SYNTHESIZED IN VITRO & MAY
CC  PLAY SOME ROLE IN DNA REPLICATION (BY SIMILARITY).
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; U46933; AAC54905.1; -.
DR  InterPro; IPR003391; Adeno terminal.
DR  Pfam; PF02459; Adeno terminal; 1.
KW  DNA replication; Covalent protein-DNA linkage.
FT  BINDING 510 510 COVALENT LINKAGE OF VIRAL DNA
FT  (BY SIMILARITY)
SQ  SEQUENCE 575 AA; 66092 MW; 0E1B86D678528437 CRC64;

Query Match 78.0%; Score 32; DB 1; Length 575;
Best Local Similarity 100.0%; Pred.No. 32;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  2 TALEAIG 8
DB  |||||

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=cv, Columbia;
RX MEDLINE=20181125; PubMed=10718197;
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
RT clones.";
RL DNA Res. 7:31-63(2000).
CC -1- FUNCTION: POSSIBLE ROLE IN MITOSIS.
CC -1- DOMAIN: COMPOSED OF THREE STRUCTURAL DOMAINS; A SMALL GLOBULAR
CC N-TERMINAL, A CENTRAL ALPHA-HELICAL COILED COIL AND A LARGE
CC GLOBULAR C-TERMINAL WHICH IS RESPONSIBLE FOR THE MOTOR ACTIVITY
CC (IT HYDROLYZES ATP AND BINDS MICROTUBULES).
CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. NCD
CC SUBFAMILY.
CC -1- CAUTION: Ref.2 sequence differs from that shown due to erroneous
CC gene model prediction.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D21138; BAA04674.1; -.
CC EMBL; AB022214; BAB09933.1; ALT_SEQ.
CC PIR; S48020; S48020.
CC HSPF; P17119; 3KAR.
CC InterPro; IPR001752; kinesin_motor.
CC Pfam; PF00225; Kinesin; 1.
CC PRINTS; PR00380; KINESINHEAVY.
CC SMART; SM00129; KISC; 1.
CC PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
CC PROSITE; PS00067; KINESIN MOTOR DOMAIN2; 1.
CC Motor protein; Microtubules; ATP-binding; Coiled coil; Mitosis;
KW Multigene family.
FT DOMAIN 1 45 GLOBULAR.
FT DOMAIN 46 393 COILED COIL.
FT DOMAIN 398 727 KINESIN-MOTOR (BY SIMILARITY).
FT NP_BIND 481 488 ATP (POTENTIAL).
SQ SEQUENCE 754 AA; 85030 MW; 76091CDB5B5D9C531 CRC64;

Query Match 78.0%; Score 32; DB 1; Length 754;
Best Local Similarity 75.0%; Pred.No. 41;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TALEAIGR 9
Db 423 TSLEALGR 430

RESULT 7
ID1 MIP1_SCHPO STANDARD; PRT; 1313 AA.
AC PB7141;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE WD-repeat protein mipl.
GN MIP1 OR SPAC57A7.11.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OC NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.; FUNCTION, AND SUBCELLULAR LOCATION.
RX MEDLINE=20115869; PubMed=10648609;
RA Shinozaki-Yabana S., Watanabe Y., Yamamoto M.;
RT "Novel WD-repeat protein Mip1 facilitates function of the meiotic

```

```

RT regulator Mei2p in fission yeast.";
RL Mol. Cell. Biol. 20:1234-1242(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones I., Jones M., Leather S., McDonald S., McLean J.,
RA Money P., Meule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Voickert G., Aert R., Robben J., Grymonprez B.,
RA Weijens J., Voickert G., Aert R., Robben J., Grymonprez B.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehtach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu B., Driano S., Gloux S., Delaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Talleza V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrett B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -1- REGULATOR BINDS TO AND FACILITATES THE FUNCTIONING OF THE MEIOTIC
CC REGULATOR MEI2. MAY ALSO BE INVOLVED IN CONJUGATION BY INTERACTING
CC WITH STELL. ESSENTIAL FOR CELL GROWTH.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Contains 7 WD repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB032552; BAA84585.1; -.
CC EMBL; Z95396; CAB08769.1; -.
CC PIR; T38943; T38943.
CC GeneDB SPombe; SPAC57A7.11; -.
CC InterPro; IPR001680; WD40.
CC InterPro; IPR004083; Yeast176.
CC Pfam; PF00400; WD40; 5.
CC PRINTS; PR01547; YEAST176DUF.
CC SMART; SM00320; WD40; 6.
CC PROSITE; PS00678; WD_REPEATS_1; 1.
CC PROSITE; PS00082; WD_REPEATS_2; 2.
CC PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Melcois; WD repeat; Repeat.
FT REPEAT 986 1029 WD 1.
FT REPEAT 1033 1074 WD 2.
FT REPEAT 1087 1126 WD 3.
FT REPEAT 1130 1170 WD 4.
FT REPEAT 1176 1216 WD 5.
FT REPEAT 1219 1259 WD 6.
FT REPEAT 1268 1308 WD 7.
SQ SEQUENCE 1313 AA; 148533 MW; C71B663B0171E7A4 CRC64;

Query Match 78.0%; Score 32; DB 1; Length 1313;
Best Local Similarity 77.8%; Pred.No. 72;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 STALEAIGR 9

```

```

RESULT 10
TOR1 SCHPO STANDARD; PRT; 2335 AA.
AC O14356;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phosphatidylinositol 3-kinase tor1 (EC 2.7.1.137) (PI3-kinase)
DE (Ptdins-3-kinase) (PI3K).
GN TOR1 OR SPBC30D10.10C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayes J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule L., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu B., Dreano S., Gloux S., Leleau V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode J.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Foreburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrill B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880(2002).
RN [2]
RP IDENTIFICATION, FUNCTION, AND INDUCTION.
RX MEDLINE=21269225; PubMed=11096119;
RA Weisman R., Choder M.;
RT "The fission yeast TOR homolog, tor1+, is required for the response to
RT starvation and other stresses via a conserved serine."
RL J. Biol. Chem. 276:7027-7032(2001).
CC -!- FUNCTION: Phosphatidylinositol 3-kinase homolog required for G1
CC progression and entry into stationary phase. Also required for the
CC onset of meiosis and sporulation under nitrogen and carbon
CC starvation conditions.
CC -!- CATALYTIC ACTIVITY: ATP + 1-phosphatidyl-1D-myo-inositol = ADP +
CC 1-phosphatidyl-1D-myo-inositol 3-phosphate.
CC -!- INDUCTION: By nitrogen and/or carbon starvation, cold, osmotic and
CC oxidative stress.
CC -!- SIMILARITY: BELONGS TO THE PI3/PI4-KINASES FAMILY.
CC -!- SIMILARITY: Contains 16 HEAT repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z97992; CAB10805.1; -.
CC F01; T40186; T40186.

```

```

DR HSP; P42345; 1FAP.
DR GeneDB SPombe; SPBC30D10.10C; -.
DR InterPro; IPR003151; FAT.
DR InterPro; IPR003152; FATC.
DR InterPro; IPR000357; HEAT_repeat.
DR InterPro; IPR000403; PI3_PI4_kinase.
DR Pfam; PF02259; FAT; 1.
DR Pfam; PF02260; FATC; 1.
DR Pfam; PF00454; PI3_PI4_kinase; 1.
DR SMART; SM00146; PI3Kc; 1.
DR PROSITE; PS50077; HEAT_REPEAT; FALSE_NEG.
DR PROSITE; PS00915; PI3_4_KINASE_1; 1.
DR PROSITE; PS00916; PI3_4_KINASE_2; 1.
DR PROSITE; PS0290; PI3_4_KINASE_3; 1.
DR Transferase; Kinase; Cell cycle; Repeat.
FT REPEAT 1 31 HEAT 1.
FT REPEAT 164 201 HEAT 2.
FT REPEAT 331 371 HEAT 3.
FT REPEAT 410 449 HEAT 4.
FT REPEAT 474 512 HEAT 5.
FT REPEAT 522 560 HEAT 6.
FT REPEAT 562 596 HEAT 7.
FT REPEAT 642 679 HEAT 8.
FT REPEAT 684 722 HEAT 9.
FT REPEAT 728 766 HEAT 10.
FT REPEAT 843 880 HEAT 11.
FT REPEAT 904 923 HEAT 12.
FT REPEAT 924 961 HEAT 13.
FT REPEAT 964 1003 HEAT 14.
FT REPEAT 1005 1042 HEAT 15.
FT REPEAT 1732 1769 HEAT 16.
FT DOMAIN 1987 2335 PI3K/PI4K.
SQ SEQUENCE 2335 AA; 266181 MW; SDFP1CF4ABE9BA4 CRC64;

Query Match 75.6%; Score 31; DB 1; Length 2335;
Best Local Similarity 77.8%; Pred. No. 2.1e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 STALEAIGR 9
DB 312 SLAFAIGR 320

RESULT 11
DEF3 RABIT STANDARD; PRT; 95 AA.
ID DEF3 RABIT STANDARD; PRT; 95 AA.
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Corticostatin III precursor (CS-III) (Macrophage antibiotic peptide
DE MCP-1) (NP-1) (Antidrenocorticotropin peptide III).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89309825; PubMed=2745983;
RA Ganz T., Rayner J.R., Valore E.V., Tumolo A., Talmadge K., Fuller F.;
RT "The structure of the rabbit macrophage defensin genes and their
RT organ-specific expression".
RL J. Immunol. 143:1358-1365(1989).
RN [2]
RP SEQUENCE OF 63-95.
RC TISSUE=Lung macrophage;
RX MEDLINE=84061901; PubMed=6643497;
RA Seasted M.E., Brown D.M., Delange R.J., Lehrer R.I.;
RT "Primary structures of MCP-1 and MCP-2, natural peptide antibiotics
RT of rabbit lung macrophages."
RL J. Biol. Chem. 258:14485-14489(1983).
RN [3]
RP SEQUENCE OF 63-95.

```



```

DR InterPro; IPR006081; Defensin_alpha.
DR InterPro; IPR006080; Defensin_mammal.
DR InterPro; IPR002366; Defensin_propep.
DR Pfam; PF00879; Defensin_propep; 1.
DR Pfam; PF00323; defensins; 1.
DR SMART; SM00048; DEFNS; 1.
DR PROSITE; PS00269; DEFENSIN; 1.
DR Defensin; Antibiotic; Antiviral; Fungicide; Signal.
FT SIGNAL 1 19
FT PROPEP 20 62
FT CHAIN 63 95
FT DISULFID 65 93
FT DISULFID 67 82
FT DISULFID 72 92
FT DISULFID 72 92
SQ SEQUENCE 95 AA; 10431 MW; 58418C82B462F332 CRC64;

Query Match 73.2%; Score 30; DB 1; Length 95;
Best Local Similarity 75.0%; Pred. No. 15;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 STALEAIG 8
DB 51 SSALFALG 58

RESULT 13
TPIS TREPA STANDARD; PRT; 249 AA.
AC 083578;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Triosephosphate isomerase (EC 5.3.1.1) (TIM).
GN TPPI OR TPI OR TP0537.
OS Treponema pallidum.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
ON NCBI_TaxID=160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nichols;
RX MEDLINE=9832770; PubMed=9658976;
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardam J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Uterback T.,
RA McDonald L., Attiash P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RA "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete.";
RL Science 281:375-388(1998).
CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate = glyceralone
phosphate.
CC -1- PATHWAY: Plays an important role in several metabolic pathways.
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SIMILARITY: BELONGS TO THE TRIOSEPHOSPHATE ISOMERASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE001229; AAC65522.1; -.
CC PIR; F71311; F71311.
CC HSP; P34204; 1B9B.
CC TIGR; TP0537; -.
CC HAWAP; MF 00147; -.
CC InterPro; IPR000652; Triophos_ismrse.
CC Pfam; PF00121; TIM; 1.
CC ProDom; PDC01005; Triophos_ismrse; 1.

```

```

DR TIGREAFMS; TIGR00419; tim; 1.
DR PROSITE; PS00171; TIM; 1.
KW Isomerase; Glycolysis; Gluconogenesis; Fatty acid biosynthesis;
KW Pentose shunt; Complete proteome.
FT ACT SITE 94 94 BY SIMILARITY.
FT ACT SITE 166 166 BY SIMILARITY.
SQ SEQUENCE 249 AA; 26540 MW; EF65741D5EE4D431 CRC64;

Query Match 73.2%; Score 30; DB 1; Length 249;
Best Local Similarity 62.5%; Pred. No. 38;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 TALEAIGR 9
DB 44 TALDAVGK 51

RESULT 14
FEN HALNI STANDARD; PRT; 327 AA.
ID FEN HALNI
AC Q9HQ27;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Flap structure-specific endonuclease (EC 3.1.1.1).
GN FEN OR VNG1359G.
OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
ON NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leithausen B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Beck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;
RA "Genome sequence of Halobacterium species NRC-1.";
RT Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
CC -1- FUNCTION: Endonuclease that cleave the 5'overhanging flap
structure that is generated by displacement synthesis when DNA
polymerase encounters the 5'end of a downstream Okazaki fragment.
CC Has 5'endo-/exonuclease and 5'pseudo-Y-endonuclease activities.
CC Cleaves the junction between single and double-stranded regions of
CC flap DNA (By similarity).
CC -1- COFACTOR: Binds 2 magnesium ions per subunit (By similarity).
CC -1- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. FEN1
SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE005056; AAG19690.1; -.
CC PIR; F84290; F84290.
CC HSP; Q58839; 1A76.
CC HAWAP; MF 00614; -.
CC InterPro; IPR000513; Exo_N_I.
CC InterPro; IPR006086; XPG_I.
CC InterPro; IPR006085; XPG_N.
CC InterPro; IPR006084; XPGC_Rad.
CC Pfam; PF00867; XPG_I; 1.
CC Pfam; PF00752; XPG_N; 1.
CC PRINTS; PR00853; XPGRADSUPER.
CC SMART; SM00484; XPGI; 1.

```

C;Superfamily: phage T7 head-to-tail joining protein

Query Match 90.2%; Score 37; DB 1; Length 536;
Best Local Similarity 88.9%; Pred. No. 3.5;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 STALEAIGR 9
||| |||||
DB 427 STGLEAIGR 435

RESULT 3

T18737

hypothetical protein B0393.3 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000

C;Accession: T18737

R;Sulston, J.

Submitted to the EMBL Data Library, September 1994

A;Reference number: Z19013

A;Accession: T18737

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-621 <Wtl>

A;Cross-references: EMBL:Z37983; PIDN:CAA86056.1; GSPDB:GN00021; CESP:B0393.3

A;Experimental source: clone B0393

C;Genetics:

A;Gene: CESP:B0393.3

A;Map position: 3

A;Introns: 31/3; 182/3; 394/3; 529/3

C;Superfamily: Caenorhabditis elegans hypothetical protein B0393.3

Query Match

80.5%; Score 33; DB 2; Length 621;

Best Local Similarity 77.6%; Pred. No. 33;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 STALEAIGR 9
||| |||||
DB 336 STGLEAIGR 344

RESULT 4

A69140

ATP-dependent helicase - Methanobacterium thermoautotrophicum (strain Delta H)

C;Species: Methanobacterium thermoautotrophicum

C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 02-Feb-2001

C;Accession: A69140

R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;

Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiواني, N.

Ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.

J. Bacteriol. 179, 7135-7155, 1997

A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: func

A;Reference number: A69000; MUID:98037514; PMID:9371463

A;Accession: A69140

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-702 <Wtl>

A;Cross-references: GB:AF000816; GB:AF000666; MID:G2621361; PIDN:AA84818.1; PID:G262136

A;Experimental source: strain Delta H

C;Genetics:

A;Gene: MTH312

C;Keywords: ATP; nucleotide binding; P-loop

F;44-51/Region: nucleotide-binding motif A (P-loop)

F;151-156/Region: nucleotide-binding motif B

F;155-158/Region: DEXH motif

Query Match

80.5%; Score 33; DB 2; Length 702;

Best Local Similarity 87.5%; Pred. No. 38;

Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 TALEAIGR 9
: |||||
DB 29 SALEAIGR 36

RESULT 5

E95963

probable transcription regulator, gntR family protein [imported] - Sinorhizobium melilo

C;Species: Sinorhizobium meliloti

C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 14-Dec-2001

C;Accession: E95963

R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Verholter, F.J.; Herna

Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001

A;Title: The complete sequence of the 1.683-kb pSymB megaplasmid from the N2-fixing end

A;Reference number: A95842; MUID:21396508; PMID:11481431

A;Accession: E95963

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-261 <Kur>

A;Cross-references: GB:AL591985; PIDN:CA49373.1; PID:g15140859; GSPDB:GN00167

A;Experimental source: strain 1021, megaplasmid pSymB

R;Galibert, F.; Finan, T.M.; Long, S.R.; Fuhler, A.; Abola, P.; Ampe, P.; Barloy-Hubler

Pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.

L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure

hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K

A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A;Reference number: A96039; MUID:21368234; PMID:11474104

A;Contents: annotation

C;Genetics:

A;Gene: SMD21533

A;Genome: plasmid

C;Superfamily: regulatory protein gntR

Query Match 78.0%; Score 32; DB 2; Length 261;

Best Local Similarity 100.0%; Pred. No. 22;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 ALEAIGR 9
||| |||||
DB 163 ALEAIGR 169

RESULT 6

AB1878

Cyanophycinase [imported] - Nostoc sp. (strain PCC 7120)

C;Species: Nostoc sp. PCC 7120

A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002

C;Accession: AB1878

R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguch

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata,

DNA Res. 8, 205-213, 2001

A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An

A;Reference number: AB1807; MUID:21595285; PMID:11759840

A;Accession: AB1878

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-293 <Kur>

A;Cross-references: GB:BA000019; PIDN:BA72529.1; PID:g17129916; GSPDB:GN00179

A;Experimental source: strain PCC 7120

C;Genetics:

A;Gene: all0571

Query Match

78.0%; Score 32; DB 2; Length 293;

Best Local Similarity 77.8%; Pred. No. 25;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 STALEAIGR 9

||| |||||

DB 88 STALEAIGR 96

RESULT 7

GB7133

```

A:Molecule type: DNA
A:Residues: 1-362 <GLA>
A:Cross-references: GB:AL592022; PIDN:CAC97952.1; PID:G16415262; GSPDB:GN00178
A:Experimental source: strain Clp11262
C:Genetics:
A:Gene: lin2726

Query Match      78.0%; Score 32; DB 2; Length 362;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3 ALEAIGR 9
Db      354 ALEAIGR 360

RESULT 12
F84068
aspartate aminotransferase BH3350 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: F84068
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512592; PMID:11058132
A:Accession: F84068
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-393 <STO>
A:Cross-references: GB:AP001518; GB:BA000004; NID:G10175792; PIDN:BA807069.1; GSPDB:GN00
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH3350
C:Superfamily: aspartate transaminase

Query Match      78.0%; Score 32; DB 2; Length 393;
Best Local Similarity 87.5%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 TALEAIGR 9
Db      379 TALEAIGR 386

RESULT 13
A82042
xanthine/uracil permease family protein VC2712 [imported] - Vibrio cholerae (strain N169
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: A82042
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Basu, S.; Qin, H.; Dragoi, I.; Sellers, B
I, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: A82042
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-480 <HEI>
A:Cross-references: GB:AE004337; GB:AE003852; NID:G9657307; PIDN:AAF95852.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC2712
A:Map position: 1
C:Superfamily: hypothetical protein b2882

Query Match      78.0%; Score 32; DB 2; Length 480;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 TALEAIG 8

```

```

Db      281 TALEAIG 287

RESULT 14
T06048
kinesin-related protein katB - Arabidopsis thaliana
N:Alternate names: protein T24A18.130
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 02-Feb-2001
C:Accession: T06048; S48019
R:Bevan, M.; Van Der Schueren, J.; Chuang, Y.J.; Vost, M.; Robben, J.; Volckaert, G.; B
submitted to the Protein Sequence Database, March 1999
A:Reference number: Z15484
A:Accession: T06048
A:Molecule type: DNA
A:Residues: 1-744 <BEV>
A:Cross-references: EMBL:AL035680; GSPDB:GN00062; ATSP:T24A18.130
A:Experimental source: cultivar Columbia; BAC clone T24A18
R:Mitsui, H.; Nakatani, K.; Yamaguchi-Shinozaki, K.; Shinozaki, K.; Nishikawa, K.; Taka
Plant Mol. Biol. 25, 865-876, 1994
A:Title: Sequencing and characterization of the kinesin-related genes katB and katC of
A:Reference number: S48019; MUID:94355659; PMID:8075402
A:Accession: S48019
A:Molecule type: mRNA
A:Residues: 1-583; S', 591, 'GK', 593-744 <MIT>
A:Cross-references: EMBL:D21137; NID:G1438841; PIDN:BA04673.1; PID:G1438842
C:Genetics:
A:Gene: ATSP:T24A18.130; katB
A:Map position: 4
A:Introns: 56/3; 105/1; 141/3; 181/3; 249/3; 308/3; 333/3; 382/3; 429/1; 465/3; 519/3;
C:Superfamily: unassigned kinesin-related proteins; kinesin motor domain homology
C:Keywords: ATP; nucleotide binding; P-loop
F:388-729/Domain: kinesin motor domain homology <KXOT>
F:472-479/Region: nucleotide-binding motif A (P-loop)

Query Match      78.0%; Score 32; DB 2; Length 744;
Best Local Similarity 75.0%; Pred. No. 68;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy      2 TALEAIGR 9
Db      414 TSLEALGR 421

RESULT 15
S48020
kinesin-related protein katC - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 14-Jul-1995 #sequence_revision 01-Dec-1995 #text_change 02-Feb-2001
C:Accession: S48020
R:Mitsui, H.; Nakatani, K.; Yamaguchi-Shinozaki, K.; Shinozaki, K.; Nishikawa, K.; Taka
Plant Mol. Biol. 25, 865-876, 1994
A:Title: Sequencing and characterization of the kinesin-related genes katB and katC of
A:Reference number: S48019; MUID:94355659; PMID:8075402
A:Accession: S48020
A:Molecule type: mRNA
A:Residues: 1-754 <MIT>
A:Cross-references: EMBL:D21138; NID:G1438843; PIDN:BA04674.1; PID:di005204; PID:G1438
C:Genetics:
A:Gene: katC
C:Superfamily: unassigned kinesin-related proteins; kinesin motor domain homology
C:Keywords: ATP; nucleotide binding; P-loop
F:397-739/Domain: kinesin motor domain homology <KXOT>
F:481-488/Region: nucleotide-binding motif A (P-loop)

Query Match      78.0%; Score 32; DB 2; Length 754;
Best Local Similarity 75.0%; Pred. No. 69;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy      2 TALEAIGR 9
Db      423 TSLEALGR 430

```